

Strong Converse Exponent for State Dependent Channels With Full State Information at the Sender and Partial State Information at the Receiver

Yasutada Oohama

University of Electro-Communications, Tokyo, Japan

Email: oohama@uec.ac.jp

Abstract—We consider the state dependent channels with full state information with at the sender and partial state information at the receiver. For this state dependent channel, the channel capacity under rate constraint on the state information at the decoder was determined by Steinberg. In this paper, we study the correct probability of decoding at rates above the capacity. We prove that when the transmission rate is above the capacity this probability goes to zero exponentially and derive an explicit lower bound of this exponent function.

Keywords—State dependent channels, strong converse theorem, exponent of correct probability of decoding

I. CODING PROBLEM FOR STATE DEPENDENT CHANNELS

We consider the classical problem of channel coding with noncausal state information at the encoder, also known as the Gel'fand-Pinsker problem. In this problem, we would like to send a uniformly distributed message over a state-dependent channel $W^n : \mathcal{X}^n \times \mathcal{S}^n$, where \mathcal{S}, \mathcal{X} and \mathcal{Y} , respectively, are the state, input and output alphabets. We assume that $\mathcal{X}, \mathcal{Y}, \mathcal{S}$ are finite sets. The state-dependent channel (SDC) we study in this paper is defined by a stationary discrete memoryless channel specified by the following stochastic matrix:

$$W \triangleq \{W(y|x, s)\}_{(s,x,y) \in \mathcal{S} \times \mathcal{X} \times \mathcal{Y}}. \quad (1)$$

Let X^n be a random variable taking values in \mathcal{X}^n . We write an element of \mathcal{X}^n as $x^n = x_1 x_2 \cdots x_n$. Suppose that X^n has a probability distribution on \mathcal{X}^n denoted by $p_{X^n} = \{p_{X^n}(x^n)\}_{x^n \in \mathcal{X}^n}$. Similar notations are adopted for other random variables. Let $Y^n \in \mathcal{Y}^n$ be a random variable obtained as the channel output by connecting X^n to the input of channel under the random state S^n . We write a conditional distribution of Y^n on \mathcal{Y}^n given X^n and S^n as

$$W^n = \{W^n(y^n|x^n, s^n)\}_{(s^n, x^n, y^n) \in \mathcal{S}^n \times \mathcal{X}^n \times \mathcal{Y}^n}.$$

Since the channel is memoryless, we have

$$W^n(y^n|x^n, s^n) = \prod_{t=1}^n W(y_t|x_t, s_t). \quad (2)$$

We assume that the state information of S^n is an output of a stationary discrete memoryless source $\{S_t\}_{t=1}^n$ specified by a probability distribution $p_S = \{p_S(s)\}_{s \in \mathcal{S}}$ on \mathcal{S} . Transmission of messages via the state dependent channel is shown in Fig. 1. The random variable K_n is a message sent to the receiver. The random variable S_n represent a random state. Under S^n , a sender transforms K_n into a transmitted sequence X^n using an

encoder function $\varphi^{(n)}$ and sends it to the receiver. In this paper we consider the case where the receiver is provided with a rate limited state information. In this case encoded data $\phi^{(n)}(S^n)$ of the random state information S^n is available at the decoder. Here $\phi^{(n)}$ is an encoder function of the state information a formal definition of which is defined by

$$\phi^{(n)} : \mathcal{S}^n \rightarrow \mathcal{M}_n = \{1, 2, \dots, |\mathcal{M}_n|\}.$$

Set $M_n = \phi^{(n)}(S^n)$. In this paper we assume that the encoder function $\varphi^{(n)}$ is a stochastic encoder. In this case, $\varphi^{(n)}$ is a stochastic matrix given by

$$\varphi^{(n)} = \{\varphi^{(n)}(x^n|k, s^n)\}_{(k, s^n, x^n) \in \mathcal{K}_n \times \mathcal{S}^n \times \mathcal{X}^n},$$

where $\varphi^{(n)}(x^n|k, s^n)$ is a conditional probability of $x^n \in \mathcal{X}^n$ given $k \in \mathcal{K}_n$ and non-causal random state $s^n \in \mathcal{S}^n$. The joint probability mass function on $\mathcal{K}_n \times \mathcal{M}_n \times \mathcal{S}^n \times \mathcal{X}^n \times \mathcal{Y}^n$ is given by

$$\begin{aligned} \Pr\{(K_n, M_n, S^n, X^n, Y^n) = (k, m, s^n, x^n, y^n)\} \\ = \frac{1}{|\mathcal{K}_n|} \varphi^{(n)}(x^n|k, s^n) p_{M_n S^n}(m, s^n) \prod_{t=1}^n W(y_t|x_t, s_t) \end{aligned}$$

where $|\mathcal{K}_n|$ is a cardinality of the set \mathcal{K}_n . The decoding function at the receiver 1 is denoted by $\psi^{(n)}$. Those functions are formally defined by

$$\psi^{(n)} : \mathcal{Y}^n \times \mathcal{M}_n \rightarrow \mathcal{K}_n.$$

The average error probability of decoding on the receiver is defined by

$$\begin{aligned} P_e^{(n)} &= P_e^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}) \\ &\triangleq \Pr\{\psi^{(n)}(Y^n, M_n) \neq K_n\} \end{aligned}$$

For $k \in \mathcal{K}_n$ and $m \in \mathcal{M}_n$, set

$$\mathcal{D}(k|m) \triangleq \{y^n : \psi_1^{(n)}(y^n) = (k, m)\}.$$

A family of sets $\{\mathcal{D}(k|m)\}_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n}$ is called the decoding region. Using the decoding region, $P_e^{(n)}$ can be written as

$$\begin{aligned} P_e^{(n)} &= \frac{1}{|\mathcal{K}_n|} \sum_{(k,l) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{(s^n, x^n, y^n) \in \mathcal{S}^n \times \mathcal{X}^n \times \mathcal{Y}^n : \\ y^n \in \mathcal{D}^c(k|m)}} \\ &\quad \times \varphi^{(n)}(x^n|k, s^n) W^n(y^n|x^n, s^n) p_{M_n, S^n}(m, s^n). \end{aligned}$$

Set

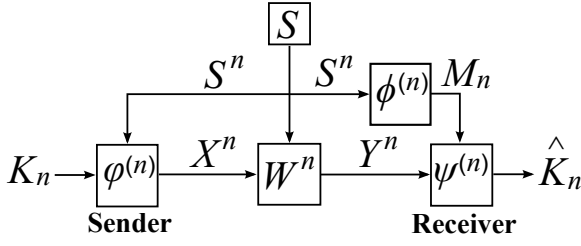


Fig. 1. Coding for state dependent channels with rate limited side Information at the receiver

$$P_c^{(n)} = P_c^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}) \triangleq 1 - P_e^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}).$$

The quantity $P_c^{(n)}$ is called the average correct probability of decoding. This quantity has the following form:

$$P_c^{(n)} = \frac{1}{|\mathcal{K}_n|} \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{(s^n, x^n, y^n) \in \mathcal{S}^n \times \mathcal{X}^n \times \mathcal{Y}^n: \\ y^n \in \mathcal{D}(k|m)}} \times \varphi^{(n)}(x^n|k, s^n) W^n(y^n|x^n, s^n) p_{M_n, S^n}(m, s^n).$$

For fixed $\varepsilon \in [0, 1)$, a pair (R_d, R) is ε -achievable if there exists a sequence of triples $\{(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)})\}_{n=1}^\infty$ such that

$$\limsup_{n \rightarrow \infty} P_e^{(n)} \leq \varepsilon, \\ \liminf_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{K}_n| \geq R, \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{M}_n| \leq R_d.$$

The set that consists of all achievable rate pair is denoted by $\mathcal{C}(\varepsilon|W)$, which is called the capacity region of the state dependent channels. Furthermore, set

$$C(\varepsilon, R_d|W) \triangleq \max_{(R_d, R) \in \mathcal{C}(\varepsilon|W)} R.$$

It is obvious that the determination problem of $\mathcal{C}(\varepsilon|W)$ is equivalent to that of $C(\varepsilon, R_d|W)$ for fixed $R_d > 0$.

To describe previous works on $\mathcal{C}(\varepsilon|W)$, we introduce a pair of auxiliary random variables (U, V) taking values in a finite set $\mathcal{U} \times \mathcal{V}$. We assume that the joint distribution of (U, V, S, X, Y) is

$$p_{UVXYZ}(u, v, s, x, y) \\ = p_{UV}(u, v) p_{SX|UV}(s, x|u, v) W(y|x, s).$$

The above condition is equivalent to $(U, V) \leftrightarrow (X, S) \leftrightarrow Y$. Define the set of probability distribution $p = p_{UVSXY}$ of $(U, V, S, X, Y) \in \mathcal{U} \times \mathcal{V} \times \mathcal{S} \times \mathcal{X} \times \mathcal{Y}$ by

$$\mathcal{P}(W) \triangleq \{p : |\mathcal{V}| \leq |\mathcal{S}||\mathcal{X}| + 1, |\mathcal{U}| \leq |\mathcal{V}||\mathcal{S}||\mathcal{X}|, \\ p_{Y|XS} = W, (U, V) \leftrightarrow (S, X) \leftrightarrow Y\}.$$

Set

$$\mathcal{C}(p) \triangleq \{(R_d, R) : R, R_d \geq 0, \\ R_d \geq I_p(V; S) - I_p(V; Y), \\ R \leq I_p(U; Y|V) - I_p(U; S|V)\}, \\ \mathcal{C}(W) = \bigcup_{p \in \mathcal{P}(W)} \mathcal{C}(p).$$

We can show that the above functions and sets satisfy the following property.

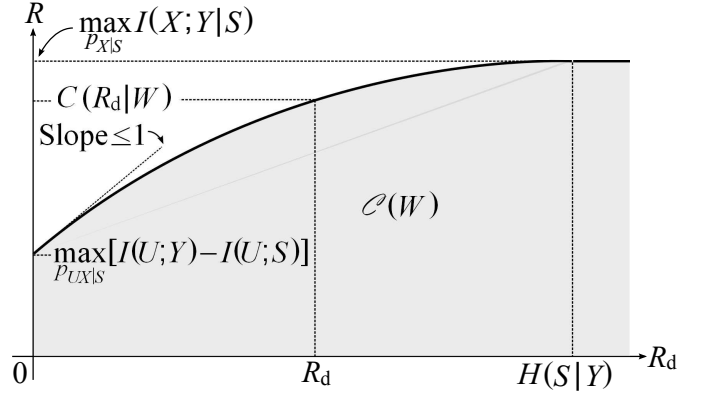


Fig. 2. Shape of $\mathcal{C}(W)$. In this figure $C(R_d|W)$ is defined by $C(R_d|W) \triangleq \max\{R : (R_d, R) \in \mathcal{C}(W)\}$.

Property 1:

- a) The region $\mathcal{C}(W)$ is a closed convex subset of \mathbb{R}_+^2 , where

$$\mathbb{R}_+^2 \triangleq \{(R_d, R) : R_d \geq 0, R \geq 0\}.$$

- b) The region $\mathcal{C}(W)$ can be expressed with two families of supporting hyperplanes. To describe this result we define the set of probability distribution $p = p_{UVSXY}$ of $(U, V, S, X, Y) \in \mathcal{U} \times \mathcal{V} \times \mathcal{S} \times \mathcal{X} \times \mathcal{Y}$ by

$$\mathcal{P}_{sh}(W) \triangleq \{p : |\mathcal{V}| \leq \min\{|\mathcal{S}||\mathcal{X}|, |\mathcal{S}| + |\mathcal{Y}| - 1\}, \\ |\mathcal{U}| \leq \min\{|\mathcal{V}||\mathcal{S}||\mathcal{X}|, |\mathcal{S}| + |\mathcal{Y}| - 1\}, \\ p_{Y|XS} = W, (U, V) \leftrightarrow (S, X) \leftrightarrow Y\}.$$

Furthermore, we define the set of probability distribution $q = q_{UVSXY}$ of $(U, V, S, X, Y) \in \mathcal{U} \times \mathcal{V} \times \mathcal{S} \times \mathcal{X} \times \mathcal{Y}$ by

$$\mathcal{Q} \triangleq \{q = q_{UVSXY} : |\mathcal{U}|, |\mathcal{V}| \leq |\mathcal{S}| + |\mathcal{Y}| - 1\}.$$

We set

$$C^{(\mu)}(W) \triangleq \max_{p \in \mathcal{P}_{sh}(W)} \{I_p(U; Y|V) - I_p(U; S|V) \\ - \mu[I_p(V; S) - I_p(V; Y)]\},$$

$$\tilde{C}^{(\alpha, \mu)}(W) \triangleq \max_{q \in \mathcal{Q}} \{-\alpha D(q_{Y|XSUV} || W | q_{XSUV}) \\ + I_q(U; Y|V) - I_q(U; S|V) \\ - \mu[I_q(V; S) - I_q(V; Y)]\},$$

$$\mathcal{C}_{sh}(W) \triangleq \bigcap_{\mu > 0} \{(R_d, R) : R - \mu R_d \leq C^{(\mu)}(W)\},$$

$$\tilde{\mathcal{C}}_{sh}(W) \triangleq \bigcap_{\alpha, \mu > 0} \{(R_d, R) : R - \mu R_d \leq \tilde{C}^{(\alpha, \mu)}(W)\}.$$

Then, we have the following:

$$\mathcal{C}(W) = \mathcal{C}_{sh}(W) = \tilde{\mathcal{C}}_{sh}(W).$$

Property 1 part a) is a well known result. Proof of Property 1 part a) is omitted. Proof of Property 1 part b) is given in the Appendix B. Typical shape of the region $\mathcal{C}(W)$ is shown in Fig. 2.

Coding problem in the case of $R_d = 0$ is called Gelfand and Pinsker problem, which was posed and investigated by Gelfand and Pinsker [1]. They determined $C(0, 0|W)$. Their result is the following:

Theorem 1 (Gelfand and Pinsker [1]): For any state dependent channel W ,

$$C(0, 0|W) = \max_{\substack{p_{U\mathcal{X}|S}: \\ |\mathcal{U}| \leq |\mathcal{S}| + 1}} \{I(U; Y) - I(U; S)\}.$$

Strong converse theorem is proved by Tyagi and Narayan [2]. Their result is the following:

Theorem 2 (Tyagi and Narayan [2]): For each $\varepsilon \in [0, 1)$, and for any state dependent channel W , we have

$$C(\varepsilon, 0|W) = C(0, 0|W).$$

To prove this theorem they used a method of image size characterization introduced by Csiszár and Körner [3].

On the determination problem of $\mathcal{C}(0|W)$ posed and investigated by Heegard and El Gamal [4], they proved that $\mathcal{C}(W)$ serves as an inner bound of the capacity region $\mathcal{C}(0|W)$. That is, we have the following:

Theorem 3 (Heegard and El Gamal [4]): For any state dependent channel W , we have

$$\mathcal{C}(0|W) \supseteq \mathcal{C}(W).$$

Subsequently, Steinberg [5] proved that the inner bound $\mathcal{C}(W)$ is tight, thereby establishing the following theorem:

Theorem 4 (Steinberg [5]): For any state dependent channel W , we have

$$\mathcal{C}(0|W) = \mathcal{C}(W).$$

In this paper we shall prove that the strong converse theorem holds for the state dependent channel with full state information and partial state information at the encoder, i.e., we have $\mathcal{C}(\varepsilon|W) = \mathcal{C}(W)$ for each $\varepsilon \in [0, 1)$.

Capacity theorems for the state dependent channel in the case of general noisy channels was obtained by Tan [6]. To derive those capacity results he used the information spectrum method introduced by Han [7].

To examine an asymptotic behavior of $P_c^{(n)}$ for rates outside the capacity region $\mathcal{C}(W)$ we define the following quantity.

$$\begin{aligned} G^{(n)}(R_d, R|W) &\triangleq \min_{\substack{(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}): \\ (1/n) \log |\mathcal{M}_n| \leq R_d, \\ (1/n) \log |\mathcal{K}_n| \geq R}} \left(-\frac{1}{n} \right) \log P_c^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}). \end{aligned}$$

By time sharing we have that $\{G^{(n)}(R_d, R|W)\}_{n \geq 1}$ satisfies the following subadditivity property:

$$\begin{aligned} &G^{(n+m)}(R_d, R|W) \\ &\leq \frac{nG^{(n)}(R_d, R|W) + mG^{(m)}(R_d, R|W)}{n+m}. \end{aligned}$$

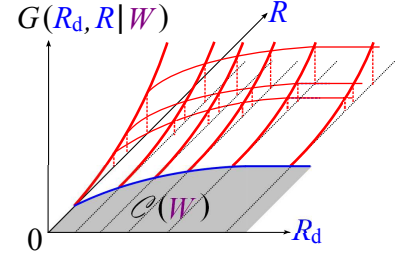


Fig. 3. Typical shape of $G(R_d, R|W)$.

Hence we have

$$\lim_{n \rightarrow \infty} G^{(n)}(R_d, R|W) = \inf_{n \geq 1} G^{(n)}(R_d, R|W).$$

Set

$$G(R_d, R|W) = \inf_{n \geq 1} G^{(n)}(R_d, R|W),$$

$$\mathcal{R}(W) \triangleq \{(R_d, R, G) : G \geq G(R_d, R|W)\}.$$

The exponent function $G(R_d, R|W)$ is a convex function of (R, R_d) . In fact, by time sharing we have that

$$\begin{aligned} &G^{(n+m)}\left(\frac{nR_d + mR'_d}{n+m}, \frac{nR + mR'}{n+m} \middle| W\right) \\ &\leq \frac{nG^{(n)}(R_d, R|W) + mG^{(m)}(R'_d, R'|W)}{n+m}, \end{aligned}$$

from which we have that for any $\alpha \in [0, 1]$

$$\begin{aligned} &G(\alpha R_d + \bar{\alpha} R'_d, \alpha R + \bar{\alpha} R'|W) \\ &\leq \alpha G(R_d, R|W) + \bar{\alpha} G(R'_d, R'|W). \end{aligned}$$

The region $\mathcal{R}(W)$ is also a closed convex set. Shape of $G(R_d, R|W)$ is shown in Fig. 3. Our main aim is to find an explicit characterization of $\mathcal{R}(W)$. In this paper we derive an explicit outer bound of $\mathcal{R}(W)$ whose section by the plane $G = 0$ coincides with $\mathcal{C}(W)$.

II. MAIN RESULT

In this section we state our main result. Define

$$\begin{aligned} &\omega_q^{(\alpha, \mu)}(s, x, y|u, v) \\ &\triangleq \alpha \log \frac{W(y|x, s)}{q_{Y|XSUV}(y|x, s, u, v)} \\ &\quad + \log \frac{q_{Y|UV}(y|u, v) q_{S|V}(s|v)}{q_{S|UV}(y|u, v) q_{Y|V}(y|v)} - \mu \log \frac{q_{S|V}(s|v) q_Y(y)}{q_{Y|V}(y|v) q_S(s)}, \\ &\Lambda_q^{(\alpha, \mu, \lambda)}(SXY|UV) \\ &\triangleq \sum_{\substack{(u, v, s, x, y) \\ \in \mathcal{U} \times \mathcal{V} \times \mathcal{S} \times \mathcal{X} \times \mathcal{Y}}} q_{UVSXY}(u, v, s, x, y) \\ &\quad \times \exp \left\{ \lambda \omega_q^{(\alpha, \mu)}(s, x, y|u, v) \right\}, \\ &\Omega_q^{(\alpha, \mu, \lambda)}(SXY|UV) \triangleq \log \Lambda_q^{(\alpha, \mu, \lambda)}(SXY|UV), \\ &\Omega^{(\alpha, \mu, \lambda)}(W) \triangleq \max_{q \in \mathcal{Q}} \Omega_q^{(\alpha, \mu, \lambda)}(SXY|UV), \\ &F^{(\alpha, \mu, \lambda)}(R_d, R|q) \end{aligned}$$

$$\begin{aligned}
&\triangleq \frac{\lambda(R - \mu R_d) - \Omega_q^{(\alpha, \mu, \lambda)}(SXY|UV)}{1 + \lambda(4 + \alpha + 3\mu)}, \\
&F^{(\alpha, \mu, \lambda)}(R_d, R|W) \\
&\triangleq \min_{q \in \mathcal{Q}} F^{(\alpha, \mu, \lambda)}(R_d, R|q) = \frac{\lambda(R - \mu R_d) - \Omega_q^{(\alpha, \mu, \lambda)}(W)}{1 + \lambda(4 + \alpha + 3\mu)}, \\
&F(R_d, R|W) \triangleq \sup_{\alpha, \mu, \lambda > 0} F^{(\alpha, \mu, \lambda)}(R_d, R|W), \\
&\overline{\mathcal{R}}(W) \triangleq \{(R_d, R, G) : G \geq F(R_d, R|W)\}.
\end{aligned}$$

We can show that the above functions and sets satisfy the following property.

Property 2:

a) $\Omega_q^{(\alpha, \mu, \lambda)}(SXY|UV)$ is a convex function of $\lambda > 0$.

b) For every $q \in \mathcal{Q}$, we have

$$\begin{aligned}
&\lim_{\lambda \rightarrow +0} \frac{\Omega_q^{(\alpha, \mu, \lambda)}(SXY|UV)}{\lambda} \\
&= -\alpha D(q_{Y|XSUUV} || W | q_{XSUV}) \\
&\quad + I_q(U; Y|V) - I_q(U; S|V) \\
&\quad - \mu [I_q(V; S) - I_q(V; Y)].
\end{aligned}$$

c) If $(R_d, R) \notin \mathcal{C}(W)$, then we have $F(R_d, R|W) > 0$.

Proof of Property 2 is given in Appendix C. Our main result is the following.

Theorem 5: For any state dependent channel W , we have

$$G(R_d, R|W) \geq F(R_d, R|W), \quad (3)$$

$$\mathcal{R}(W) \subseteq \overline{\mathcal{R}}(W). \quad (4)$$

Proof of this theorem will be given in Section III. It follows from Theorem 5 and Property 2 part c) that if (R_d, R) is outside the capacity region, then the error probability of decoding goes to one exponentially and its exponent is not below $F(R_d, R|W)$.

From this theorem we immediately follows from the following corollary:

Corollary 1: For each $\varepsilon \in [0, 1)$, we have

$$\mathcal{C}(\varepsilon|W) = \mathcal{C}(0|W).$$

Outline of the proof of Theorem 5 will be given in the next section. The exponent function at rates outside the channel capacity was derived by Arimoto [8] and Dueck and Körner [9]. The techniques used by them are not useful to prove Theorem 5. Some novel techniques based on the information spectrum method introduced by Han [7] are necessary to prove this theorem.

III. PROOF OF THE MAIN RESULTS

We first prove the following lemma.

Lemma 1: For any $\eta > 0$ and for any $(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)})$ satisfying

$$\frac{1}{n} \log |\mathcal{K}_n| \geq R, \frac{1}{n} \log |\mathcal{M}_n| \leq R_d$$

we have

$$\begin{aligned}
&P_c^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}) \leq p_{K_n M_n S^n X^n Y^n} \left\{ \right. \\
&0 \leq \frac{1}{n} \log \frac{W^n(Y^n | X^n, S^n)}{q_{Y^n | X^n S^n K_n M_n}^{(i)}(Y^n | X^n, S^n, K_n, M_n)} + \eta, \quad (5)
\end{aligned}$$

$$0 \leq \frac{1}{n} \log \frac{p_{S^n | K_n M_n}^{(ii)}(S^n | K_n, M_n)}{q_{S^n | K_n M_n}^{(ii)}(S^n | K_n, M_n)} + \eta, \quad (6)$$

$$R \leq \frac{1}{n} \log \frac{p_{Y^n | K_n M_n}(Y^n | K_n, M_n)}{q_{Y^n | M_n}^{(iii)}(Y^n | M_n)} + \eta, \quad (7)$$

$$R_d \geq \frac{1}{n} \log \frac{q_{S^n | M_n}^{(iv)}(S^n | M_n)}{p_{S^n}(S^n)} - \eta \Big\} + 4e^{-n\eta}. \quad (8)$$

In (5), we can choose any conditional distribution $q_{Y^n | X^n}^{(i)}$ on \mathcal{Y}^n given (X^n, S^n, K_n, M_n) . In (6), we can choose any conditional distribution $q_{S^n | K_n, M_n}^{(ii)}$ on \mathcal{S}^n given (K_n, M_n) . In (7), we can choose any conditional distribution $q_{Y^n | M_n}^{(iii)}$ on \mathcal{Y}^n given M_n . In (8) we can choose any distribution $q_{S^n | M_n}^{(iv)}$ on \mathcal{S}^n given M_n .

Proof of this lemma is given in Appendix D. From this lemma we immediately obtain the following lemma.

Lemma 2: For any $\eta > 0$ and for any $(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)})$ satisfying

$$\frac{1}{n} \log |\mathcal{K}_n| \geq R, \frac{1}{n} \log |\mathcal{M}_n| \leq R_d,$$

we have

$$\begin{aligned}
&P_c^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}) \leq p_{K_n M_n S^n X^n Y^n} \left\{ \right. \\
&0 \leq \frac{1}{n} \log \frac{W^n(Y^n | X^n, S^n)}{q_{Y^n | X^n S^n K_n M_n}^{(i)}(Y^n | X^n, S^n, K_n, M_n)} + \eta, \quad (9)
\end{aligned}$$

$$R \leq \frac{1}{n} \log \frac{q_{Y^n | K_n M_n}^{(ii)}(Y^n | K_n, M_n) q_{S^n | M_n}^{(iii)}(S^n | M_n)}{q_{S^n | K_n M_n}^{(ii)}(S^n | K_n, M_n) q_{Y^n | M_n}^{(iii)}(Y^n | M_n)} \quad (10)$$

$$+ \frac{1}{n} \log \frac{p_{Y^n | K_n M_n}(Y^n | K_n, M_n)}{q_{Y^n | K_n M_n}^{(ii)}(Y^n | K_n, M_n)} \quad (11)$$

$$+ \frac{1}{n} \log \frac{p_{S^n | K_n M_n}(S^n | K_n, M_n)}{q_{S^n | M_n}^{(iii)}(S^n | M_n)} + 2\eta, \quad (12)$$

$$R_d \geq \frac{1}{n} \log \frac{q_{S^n | M_n}^{(iv)}(S^n | M_n) q_{Y^n}^{(v)}(Y^n)}{q_{Y^n | M_n}^{(iv)}(Y^n | M_n) q_{S^n}^{(vi)}(S^n)} \quad (13)$$

$$+ \frac{1}{n} \log \frac{q_{S^n}^{(vi)}(S^n) q_{Y^n | M_n}^{(iv)}(Y^n | M_n)}{p_{S^n}(S^n) q_{Y^n}^{(v)}(Y^n)} - \eta \Big\} + 4e^{-n\eta}. \quad (14)$$

In (9), the choice of $q_{Y^n | X^n S^n K_n M_n}^{(i)}$ is the same as (5) in Lemma 1. In (10) and (11), we can choose any pair of $(q_{Y^n | K_n M_n}^{(ii)}, q_{S^n | K_n M_n}^{(ii)})$. In (10) and (12), we can choose any pair of $(q_{Y^n | M_n}^{(iii)}, q_{S^n | M_n}^{(iii)})$. In (13) we can choose any pair of $(q_{Y^n | M_n}^{(iv)}, q_{S^n | M_n}^{(iv)})$. In (13) and (14), we can choose any

distribution $q_{Y^n}^{(v)}$ on \mathcal{Y}^n . In (13) and (14), we can choose any distribution $q_{S^n}^{(vi)}$ on \mathcal{S}^n .

Proof: From Lemma 1, we have the following chain of inequalities:

$$\begin{aligned}
P_c^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}) &\leq p_{K_n M_n S^n X^n Y^n} \left\{ \right. \\
0 &\leq \frac{1}{n} \log \frac{W^n(Y^n|X^n, S^n)}{q_{Y^n|X^n S^n K_n M_n}^{(i)}(Y^n|X^n, S^n, K_n, M_n)} + \eta, \\
R &\leq \frac{1}{n} \log \frac{p_{Y^n|K_n M_n}(Y^n|K_n, M_n) p_{S^n|M_n}(S^n|M_n)}{q_{S^n|K_n M_n}^{(ii)}(S^n|K_n, M_n) q_{Y^n|M_n}^{(iii)}(Y^n|M_n)} \\
&\quad + 2\eta, \\
R_d &\geq \frac{1}{n} \log \frac{q_{S^n|M_n}^{(iv)}(S^n|M_n)}{p_{S^n}(S^n)} - \eta \left\} + 4e^{-n\eta} \\
&= p_{K_n M_n S^n X^n Y^n} \left\{ \right. \\
0 &\leq \frac{1}{n} \log \frac{W^n(Y^n|X^n, S^n)}{q_{Y^n|X^n S^n K_n M_n}^{(i)}(Y^n|X^n, S^n, K_n, M_n)} + \eta, \\
R &\leq \frac{1}{n} \log \frac{q_{Y^n|K_n M_n}^{(ii)}(Y^n|K_n, M_n) q_{S^n|M_n}^{(iii)}(S^n|M_n)}{q_{S^n|K_n M_n}^{(ii)}(S^n|K_n, M_n) q_{Y^n|M_n}^{(iii)}(Y^n|M_n)} \\
&\quad + \frac{1}{n} \log \frac{p_{Y^n|K_n M_n}(Y^n|K_n, M_n)}{q_{Y^n|K_n M_n}^{(ii)}(Y^n|K_n, M_n)} \\
&\quad + \frac{1}{n} \log \frac{p_{S^n|K_n M_n}(S^n|K_n, M_n)}{q_{S^n|M_n}^{(iii)}(S^n|M_n)} + 2\eta, \\
R_d &\geq \frac{1}{n} \log \frac{q_{S^n|M_n}^{(iv)}(S^n|M_n) q_{Y^n}^{(v)}(Y^n)}{q_{Y^n|M_n}^{(iv)}(Y^n|M_n) q_{S^n}^{(vi)}(S^n)} \\
&\quad + \frac{1}{n} \log \frac{q_{S^n}^{(vi)}(S^n) q_{Y^n|M_n}^{(iv)}(Y^n|M_n)}{p_{S^n}(S^n) q_{Y^n}^{(v)}(Y^n)} - \eta \left\} + 4e^{-n\eta},
\end{aligned}$$

completing the proof. \blacksquare

For $t = 1, 2, \dots, n$, set

$$\begin{aligned}
\mathcal{U}_t &\triangleq \mathcal{K}_n, U_t \triangleq K_n \in \mathcal{U}_t, \\
\mathcal{V}_t &\triangleq \mathcal{M}_n \times \mathcal{Y}^{t-1} \times \mathcal{S}_{t+1}^n, V_t \triangleq (M_n, Y^{t-1}, S_{t+1}^n) \in \mathcal{V}_t, \\
v_t &\triangleq (m, y^{t-1}, s_{t+1}^n) \in \mathcal{V}_t, \\
\hat{\mathcal{V}}_t &\triangleq \mathcal{M}_n \times \mathcal{Y}^{t-1}, \hat{V}_t \triangleq (M_n, Y^{t-1}) \in \hat{\mathcal{V}}_t, \\
\hat{v}_t &\triangleq (m, y^{t-1}) \in \hat{\mathcal{V}}_t, \\
\check{\mathcal{V}}_t &\triangleq \mathcal{M}_n \times \mathcal{S}_{t+1}^n, \check{V}_t \triangleq (M_n, S_{t+1}^n) \in \check{\mathcal{V}}_t, \\
\check{v}_t &\triangleq (m, s_{t+1}^n) \in \check{\mathcal{V}}_t.
\end{aligned}$$

From Lemma 2, we have the following.

Lemma 3: For any $\eta > 0$ and for any $(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)})$ satisfying

$$\frac{1}{n} \log |\mathcal{K}_n| \geq R, \frac{1}{n} \log |\mathcal{M}_n| \leq R_d,$$

we have

$$\begin{aligned}
P_c^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}) &\leq p_{K_n M_n S^n X^n Y^n} \left\{ \right. \\
0 &\leq \frac{1}{n} \sum_{t=1}^n \log \frac{W(Y_t|X_t, S_t)}{q_{Y_t|X_t S_t U_t V_t}^{(i)}(Y_t|X_t, S_t, U_t, V_t)} + \eta, \\
R &\leq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{Y_t|U_t V_t}^{(ii)}(Y_t|U_t, V_t) q_{S_t|V_t}^{(iii)}(S_t|V_t)}{q_{S_t|U_t V_t}^{(ii)}(S_t|U_t, V_t) q_{Y_t|V_t}^{(iii)}(Y_t|V_t)} \\
&\quad + \frac{1}{n} \sum_{t=1}^n \log \frac{p_{Y_t|U_t \hat{V}_t}(Y_t|U_t \hat{V}_t)}{q_{Y_t|U_t \hat{V}_t}^{(ii)}(Y_t|U_t \hat{V}_t)} \\
&\quad + \frac{1}{n} \sum_{t=1}^n \log \frac{p_{S_t|\check{V}_t}(S_t|\check{V}_t)}{q_{S_t|\check{V}_t}^{(iii)}(S_t|\check{V}_t)} + 2\eta, \\
R_d &\geq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{S_t|V_t}^{(iv)}(S_t|V_t) q_{Y_t}^{(v)}(Y_t)}{q_{Y_t|V_t}^{(iv)}(Y_t|V_t) q_{S_t}^{(vi)}(S_t)} \\
&\quad + \frac{1}{n} \sum_{t=1}^n \log \frac{q_{S_t}^{(vi)}(S_t) q_{Y_t|\hat{V}_t}^{(iv)}(Y_t|\hat{V}_t)}{p_{S_t}(S_t) q_{Y_t}^{(v)}(Y_t)} - \eta \left\} + 4e^{-n\eta}, \quad (15)
\end{aligned}$$

where for each $t = 1, 2, \dots, n$, the following probability and conditional probability distributions:

$$\left. \begin{aligned}
&q_{Y_t|X_t S_t U_t V_t}^{(i)}, q_{Y_t|U_t V_t}^{(ii)}, q_{Y_t|U_t \hat{V}_t}^{(ii)}, q_{S_t|U_t V_t}^{(iii)}, \\
&q_{Y_t|V_t}^{(iii)}, q_{S_t|V_t}^{(iii)}, q_{Y_t|\check{V}_t}^{(iv)}, q_{Y_t|\hat{V}_t}^{(iv)}, q_{S_t|V_t}^{(iv)}, \\
&q_{Y_t}^{(v)}, q_{S_t}^{(vi)}
\end{aligned} \right\} \quad (16)$$

appearing in the first term in the right members of (15) have a property that we can choose their values arbitrary.

Proof: On the probability distributions appearing in the right members of (14), we take the following choices. In (9), we choose $q_{Y^n|X^n S^n K_n M_n}$ so that

$$\begin{aligned}
&q_{Y^n|X^n S^n K_n M_n}^{(i)}(Y^n|X^n, S^n, K_n, M_n) \\
&= \prod_{t=1}^n q_{Y_t|X^t Y^{t-1} S_{t+1}^n K_n M_n}^{(i)}(Y_t|X_t, Y^{t-1}, S_{t+1}^n, K_n, M_n) \\
&= \prod_{t=1}^n q_{Y_t|X^t U_t V_t}^{(i)}(Y_t|X_t, S_t, U_t, V_t). \quad (17)
\end{aligned}$$

In (10), we have the following identity:

$$\begin{aligned}
&\frac{q_{Y^n|K_n M_n}^{(ii)}(Y^n|K_n, M_n)}{q_{S^n|K_n M_n}^{(ii)}(S^n|K_n, M_n)} \\
&= \prod_{t=1}^n \frac{q_{Y_t|Y^{t-1} S_{t+1}^n K_n M_n}^{(ii)}(Y_t|Y^{t-1}, S_{t+1}^n, K_n, M_n)}{q_{S_t|Y^{t-1} S_{t+1}^n K_n M_n}^{(ii)}(S_t|Y^{t-1}, S_{t+1}^n, K_n, M_n)} \\
&= \prod_{t=1}^n \frac{q_{Y_t|U_t V_t}^{(ii)}(Y_t|U_t, V_t)}{q_{S_t|U_t V_t}^{(ii)}(S_t|U_t, V_t)}. \quad (18) \\
&\frac{q_{Y^n|M_n}^{(iii)}(Y^n|M_n)}{q_{S^n|M_n}^{(iii)}(S^n|M_n)}
\end{aligned}$$

$$\begin{aligned}
&= \prod_{t=1}^n \frac{q_{Y_t|Y^{t-1}S_{t+1}^nM_n}^{(iii)}(Y_t|Y^{t-1}, S_{t+1}^n, M_n)}{q_{S_t|Y^{t-1}S_{t+1}^nM_n}^{(iii)}(S_t|Y^{t-1}, S_{t+1}^n, M_n)} \\
&= \prod_{t=1}^n \frac{q_{Y_t|V_t}^{(iii)}(Y_t|V_t)}{q_{S_t|V_t}^{(iii)}(S_t|V_t)}.
\end{aligned} \tag{19}$$

In (13), we have the following identity:

$$\begin{aligned}
&\frac{q_{Y^n|M_n}^{(iv)}(Y^n|M_n)}{q_{S^n|M_n}^{(iv)}(S^n|M_n)} \\
&= \prod_{t=1}^n \frac{q_{Y_t|Y^{t-1}S_{t+1}^nM_n}^{(iv)}(Y_t|Y^{t-1}, S_{t+1}^n, M_n)}{q_{S_t|Y^{t-1}S_{t+1}^nM_n}^{(iv)}(S_t|Y^{t-1}, S_{t+1}^n, M_n)} \\
&= \prod_{t=1}^n \frac{q_{Y_t|V_t}^{(iv)}(Y_t|V_t)}{q_{S_t|V_t}^{(iv)}(S_t|V_t)}.
\end{aligned} \tag{20}$$

In (11), we have the following:

$$\begin{aligned}
&q_{Y^n|K_nM_n}^{(ii)}(Y^n|K_n, M_n) \\
&= \prod_{t=1}^n q_{Y_t|Y^{t-1}K_nM_n}^{(ii)}(Y_t|Y^{t-1}, K_n, M_n) \\
&= \prod_{t=1}^n q_{Y_t|U_t\hat{V}_t}^{(ii)}(Y_t|U_t, \hat{V}_t).
\end{aligned} \tag{21}$$

In (12), we have the following:

$$\begin{aligned}
&\frac{p_{S^n|K_nM_n}(S^n|K_n, M_n)}{q_{S^n|M_n}^{(iii)}(S^n|M_n)} \\
&\stackrel{(a)}{=} \frac{p_{S^n|M_n}(S^n|M_n)}{q_{S^n|M_n}^{(iii)}(S^n|M_n)} = \prod_{t=1}^n \frac{p_{S_t|\tilde{V}_t}(S_t|\tilde{V}_t)}{q_{S_t|\tilde{V}_t}^{(iii)}(S_t|\tilde{V}_t)}.
\end{aligned} \tag{22}$$

Step (a) follows from that by the problem set up the state S^n is independent of K_n . In (13) and (14), we choose $q_{Y^n}^{(v)}$ and $q_{S^n}^{(vi)}$ so that

$$q_{Y^n}^{(v)}(Y^n) = \prod_{t=1}^n q_{Y_t}^{(v)}(Y_t), q_{S^n}^{(vi)}(S^n) = \prod_{t=1}^n q_{S_t}^{(vi)}(S_t). \tag{23}$$

From Lemma 2 and (17)-(23), we have

$$\begin{aligned}
&P_c^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}) \leq p_{K_nM_nS^nX^nY^n} \left\{ \right. \\
&0 \leq \frac{1}{n} \sum_{t=1}^n \log \frac{W(Y_t|X_t, S_t)}{q_{Y_t|X_tS_tU_tV_t}^{(i)}(Y_t|X_t, S_t, U_t, V_t)} + \eta, \\
&R \leq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{Y_t|U_tV_t}^{(ii)}(Y_t|U_t, V_t) q_{S_t|V_t}^{(iii)}(S_t|V_t)}{q_{S_t|U_tV_t}^{(ii)}(S_t|U_t, V_t) q_{Y_t|V_t}^{(iii)}(Y_t|V_t)} \\
&\quad + \frac{1}{n} \sum_{t=1}^n \log \frac{p_{Y_t|U_t\hat{V}_t}(Y_t|U_t, \hat{V}_t)}{q_{Y_t|U_t\hat{V}_t}^{(ii)}(Y_t|U_t, \hat{V}_t)} \\
&\quad + \frac{1}{n} \sum_{t=1}^n \log \frac{p_{S_t|\tilde{V}_t}(S_t|\tilde{V}_t)}{q_{S_t|\tilde{V}_t}^{(iii)}(S_t|\tilde{V}_t)} + 2\eta,
\end{aligned}$$

$$\begin{aligned}
R_d &\geq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{S_t|V_t}^{(iv)}(S_t|V_t) q_{Y_t}^{(v)}(Y_t)}{q_{Y_t|V_t}^{(iv)}(Y_t|V_t) q_{S_t}^{(vi)}(S_t)} \\
&\quad + \frac{1}{n} \sum_{t=1}^n \log \frac{q_{S_t}^{(vi)}(S_t) q_{Y_t|\hat{V}_t}^{(iv)}(Y_t|\hat{V}_t)}{p_{S_t}(S_t) q_{Y_t}^{(v)}(Y_t)} - \eta \Big\} + 4e^{-n\eta},
\end{aligned}$$

completing the proof. \blacksquare

For each $t = 1, 2, \dots, n$, let $\mathcal{Q}(\mathcal{U}_t \times \mathcal{V}_t \times \mathcal{S} \times \mathcal{X} \times \mathcal{Y})$ be a set of all probability distributions on

$$\mathcal{U}_t \times \mathcal{V}_t \times \mathcal{S} \times \mathcal{X} \times \mathcal{Y} = \mathcal{K}_n \times \mathcal{M}_n \times \mathcal{S}^{n-t+1} \times \mathcal{X} \times \mathcal{Y}^t$$

For $t = 1, 2, \dots, n$, we simply write $\mathcal{Q}_t = \mathcal{Q}(\mathcal{U}_t \times \mathcal{V}_t \times \mathcal{S} \times \mathcal{X} \times \mathcal{Y})$. Similarly, for $t = 1, 2, \dots, n$, we simply write $q_t = q_{U_tV_tS_tX_tY_t} \in \mathcal{Q}_t$. Set

$$\begin{aligned}
\mathcal{Q}^n &\triangleq \prod_{t=1}^n \mathcal{Q}_t = \prod_{t=1}^n \mathcal{Q}(\mathcal{U}_t \times \mathcal{V}_t \times \mathcal{S} \times \mathcal{X} \times \mathcal{Y}), \\
q^n &\triangleq \{q_t\}_{t=1}^n \in \mathcal{Q}^n.
\end{aligned}$$

From Lemma 3, we immediately obtain the following lemma.

Lemma 4: For any $\eta > 0$ and for any $(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)})$ satisfying

$$\frac{1}{n} \log |\mathcal{K}_n| \geq R, \frac{1}{n} \log |\mathcal{M}_n| \leq R_d,$$

we have

$$\begin{aligned}
&P_c^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}) \leq p_{K_nM_nS^nX^nY^n} \left\{ \right. \\
&0 \leq \frac{1}{n} \sum_{t=1}^n \log \frac{W(Y_t|X_t, S_t)}{q_{Y_t|X_tS_tU_tV_t}(Y_t|X_t, S_t, U_t, V_t)} + \eta, \\
&R \leq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{Y_t|U_tV_t}(Y_t|U_t, V_t) q_{S_t|V_t}(S_t|V_t)}{q_{S_t|U_tV_t}(S_t|U_t, V_t) q_{Y_t|V_t}(Y_t|V_t)} \\
&\quad + \frac{1}{n} \sum_{t=1}^n \log \frac{p_{Y_t|U_t\hat{V}_t}(Y_t|U_t, \hat{V}_t)}{q_{Y_t|U_t\hat{V}_t}(Y_t|U_t, \hat{V}_t)} \\
&\quad + \frac{1}{n} \sum_{t=1}^n \log \frac{p_{S_t|\tilde{V}_t}(S_t|\tilde{V}_t)}{q_{S_t|\tilde{V}_t}(S_t|\tilde{V}_t)} + 2\eta, \\
&R_d \geq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{S_t|V_t}(S_t|V_t) q_{Y_t}(Y_t)}{q_{Y_t|V_t}(Y_t|V_t) q_{S_t}(S_t)} \\
&\quad + \frac{1}{n} \sum_{t=1}^n \log \frac{q_{S_t}(S_t) q_{Y_t|\hat{V}_t}(Y_t|\hat{V}_t)}{p_{S_t}(S_t) q_{Y_t}(Y_t)} - \eta \Big\} + 4e^{-n\eta}, \tag{24}
\end{aligned}$$

where for each $t = 1, 2, \dots, n$, the following probability and conditional probability distributions:

$$\left\{ \begin{aligned} &q_{Y_t|X_tS_tU_tV_t}, q_{Y_t|U_t\hat{V}_t}, \\ &q_{Y_t|U_tV_t}, q_{Y_t|V_t}, q_{Y_t|\hat{V}_t}, q_{Y_t}, \\ &q_{S_t|U_tV_t}, q_{S_t|V_t}, q_{S_t|\tilde{V}_t}, q_{S_t} \end{aligned} \right\} \tag{25}$$

appearing in the first term in the right members of (24) are chosen so that they are induced by the joint distribution $q_t = q_{U_tV_tS_tX_tY_t} \in \mathcal{Q}_t$.

To evaluate an upper bound of (24) in Lemma 4. We use the following lemma, which is well known as the Cramér's bound in the large deviation principle.

Lemma 5: For any real valued random variable A and any $\theta > 0$, we have

$$\Pr\{A \geq a\} \leq \exp[-(\theta a - \log E[\exp(\theta A)])].$$

Here we define a quantity which serves as an exponential upper bound of $P_c^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)})$. Let $\mathcal{P}^{(n)}(W)$ be a set of all probability distributions $p_{K_n M_n S^n X^n Y^n}$ on $\mathcal{K}_n \times \mathcal{M}_n \times \mathcal{S}^n \times \mathcal{X}^n \times \mathcal{Y}^n$ having the form:

$$p_{K_n M_n S^n X^n Y^n}(k, m, s^n, x^n, y^n) = p_{K_n}(k) p_{M_n S^n}(m, s^n) \times \prod_{t=1}^n p_{X_t|K_n X^{t-1} S^n}(x_t|k, x^{t-1}, s^n) W(y_t|x_t, s_t).$$

For simplicity of notation we use the notation $p^{(n)}$ for $p_{K_n M_n S^n X^n Y^n} \in \mathcal{P}^{(n)}(W)$. We assume that $p_{U_t V_t S_t X_t Y_t} = p_{K_n M_n S_t^n X_t^n Y_t^n}$ is a marginal distribution induced by $p^{(n)}$. For $t = 1, 2, \dots, n$, we simply write $p_t = p_{U_t V_t S_t X_t Y_t}$. For $p^{(n)} \in \mathcal{P}^{(n)}(W)$ and $q^n \in \mathcal{Q}^n$, we define

$$\begin{aligned} & \Omega_{p^{(n)}||q^n}^{(\alpha, \mu, \theta)}(S^n X^n Y^n | K_n M_n) \\ & \triangleq \log E_{p^{(n)}} \left[\left\{ \prod_{t=1}^n \frac{W^{\alpha\theta}(Y_t|X_t)}{q_{Y_t|X_t S_t U_t V_t}^{\alpha\theta}(Y_t|X_t, S_t, U_t, V_t)} \right\} \right. \\ & \quad \times \left\{ \prod_{t=1}^n \frac{q_{Y_t|U_t V_t}^{\theta}(Y_t|U_t, V_t) q_{S_t|V_t}^{\theta}(S_t|V_t)}{q_{S_t|U_t V_t}^{\theta}(Y_t|U_t, V_t) q_{Y_t|V_t}^{\theta}(Y_t|V_t)} \right\} \\ & \quad \times \left\{ \prod_{t=1}^n \frac{p_{Y_t|U_t \hat{V}_t}^{\theta}(Y_t|U_t, \hat{V}_t)}{q_{Y_t|U_t \hat{V}_t}^{\theta}(Y_t|U_t, \hat{V}_t)} \right\} \left\{ \prod_{t=1}^n \frac{p_{S_t|\tilde{V}_t}^{\theta}(S_t|\tilde{V}_t)}{q_{S_t|\tilde{V}_t}^{\theta}(S_t|\tilde{V}_t)} \right\} \\ & \quad \times \left\{ \prod_{t=1}^n \frac{q_{Y_t|V_t}^{\mu\theta}(Y_t|V_t) q_{S_t}^{\mu\theta}(S_t)}{q_{S_t|V_t}^{\mu\theta}(S_t|V_t) q_{Y_t}^{\mu\theta}(Y_t)} \right\} \\ & \quad \times \left\{ \prod_{t=1}^n \frac{p_{S_t}^{\theta}(S_t)}{q_{S_t}^{\theta}(S_t)} \right\} \left\{ \prod_{t=1}^n \frac{q_{Y_t}^{\theta}(Y_t)}{q_{Y_t|\hat{V}_t}^{\theta}(Y_t|\hat{V}_t)} \right\} \Bigg], \end{aligned}$$

where for each $t = 1, 2, \dots, n$, the following probability and conditional probability distributions:

$$\left\{ \begin{aligned} & q_{Y_t|X_t S_t U_t V_t}, q_{Y_t|U_t \hat{V}_t}, \\ & q_{Y_t|U_t V_t}, q_{Y_t|V_t}, q_{Y_t|\hat{V}_t}, q_{Y_t}, \\ & q_{S_t|U_t V_t}, q_{S_t|V_t}, q_{S_t|\tilde{V}_t}, q_{S_t} \end{aligned} \right\} \quad (26)$$

appearing in the definition of $\Omega_{p^{(n)}||q^n}^{(\alpha, \mu, \theta)}(S^n X^n Y^n | K_n M_n)$ are chosen so that they are induced by the joint distribution $q_t = q_{U_t V_t S_t X_t Y_t} \in \mathcal{Q}_t$. Here we give a remark on an essential difference between $p^{(n)} \in \mathcal{P}^{(n)}(W)$ and $q^n \in \mathcal{Q}^n$. For the former the n probability distributions p_t , $t = 1, 2, \dots, n$, are consistent with $p^{(n)}$, since all of them are marginal distributions of $p^{(n)}$. On the other hand, for the latter, q^n is just a sequence of n probability distributions. Hence, we may not have the consistency between the n elements q_t , $t = 1, 2, \dots, n$, of q^n .

By Lemmas 4 and 5, we have the following proposition.

Proposition 1: For any $\alpha, \mu, \theta > 0$, any $q^n \in \mathcal{Q}^n$, and any $(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)})$ satisfying

$$\frac{1}{n} \log |\mathcal{K}_n| \geq R, \frac{1}{n} \log |\mathcal{M}_n| \leq R_d,$$

we have

$$\begin{aligned} & P_c^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}) \\ & \leq 5 \exp \left\{ -n [1 + \theta(2 + \alpha + \mu)]^{-1} \right. \\ & \quad \times \left. \left[\theta(R - \mu R_d) - \frac{1}{n} \Omega_{p^{(n)}||q^n}^{(\alpha, \mu, \theta)}(S^n X^n Y^n | K_n M_n) \right] \right\}. \end{aligned}$$

Proof: We define three random variables $A_i, i = 1, 2, 3$ by

$$\begin{aligned} A_1 & \triangleq \frac{1}{n} \sum_{t=1}^n \log \frac{W(Y_t|X_t S_t)}{q_{Y_t|X_t S_t U_t V_t}(Y_t|X_t, S_t, U_t, V_t)}, \\ A_2 & \triangleq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{Y_t|U_t V_t}(Y_t|U_t, V_t) q_{S_t|V_t}(S_t|V_t)}{q_{S_t|U_t V_t}(S_t|U_t, V_t) q_{Y_t|V_t}(Y_t|V_t)} \\ & \quad + \frac{1}{n} \sum_{t=1}^n \log \frac{p_{Y_t|U_t \hat{V}_t}(Y_t|U_t, \hat{V}_t)}{q_{Y_t|U_t \hat{V}_t}(Y_t|U_t, \hat{V}_t)} \\ & \quad + \frac{1}{n} \sum_{t=1}^n \log \frac{p_{S_t|\tilde{V}_t}(S_t|\tilde{V}_t)}{q_{S_t|\tilde{V}_t}(S_t|\tilde{V}_t)} - R, \\ A_3 & \triangleq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{Y_t|V_t}(Y_t|V_t) q_{S_t}(S_t)}{q_{S_t|V_t}(S_t|V_t) q_{Y_t}(Y_t)} \\ & \quad + \frac{1}{n} \sum_{t=1}^n \log \frac{p_{S_t}(S_t) q_{Y_t}(Y_t)}{q_{S_t}(S_t) q_{Y_t|\hat{V}_t}(Y_t|\hat{V}_t)} + R_d. \end{aligned}$$

Then by Lemma 4, for any $(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)})$ satisfying

$$\frac{1}{n} \log |\mathcal{K}_n| \geq R, \frac{1}{n} \log |\mathcal{M}_n| \leq R_d,$$

we have

$$\begin{aligned} & P_c^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}) \\ & = p_{K_n M_n S^n X^n Y^n} \{A_1 \geq -\eta, A_2 \geq -2\eta, A_3 \geq -\eta\} \\ & \quad + 4e^{-n\eta} \\ & \leq p_{K_n M_n S^n X^n Y^n} \{A \geq a\} + 4e^{-n\eta}, \end{aligned} \quad (27)$$

where we set

$$A \triangleq \alpha A_1 + A_2 + \mu A_3, a \triangleq -\eta[2 + \alpha + \mu].$$

Applying Lemma 5 to the first term in the right member of (27), we have

$$\begin{aligned} & P_c^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}) \\ & \leq \exp[-(n\theta a - \log E_{p^{(n)}}[\exp(n\theta A)])] + 4e^{-n\eta} \\ & = \exp \left[n \left\{ \theta(2 + \alpha + \mu)\eta - \theta[R - \mu R_d] \right. \right. \\ & \quad \left. \left. + \frac{1}{n} \Omega_{p^{(n)}||q^n}^{(\alpha, \mu, \theta)}(S^n X^n Y^n | K_n M_n) \right\} \right] + 4e^{-n\eta}. \end{aligned} \quad (28)$$

We choose η so that

$$\begin{aligned} & -\eta = \theta(2 + \alpha + \mu)\eta - \theta[R - \mu R_d] \\ & \quad + \frac{1}{n} \Omega_{p^{(n)}||q^n}^{(\alpha, \mu, \theta)}(S^n X^n Y^n | K_n M_n). \end{aligned} \quad (29)$$

Solving (29) with respect to η , we have

$$\eta = \frac{\theta[R - \mu R_d] - \frac{1}{n} \Omega_{p^{(n)}||q^n}^{(\alpha, \mu, \theta)}(S^n X^n Y^n | K_n M_n)}{1 + \theta(2 + \alpha + \mu)}.$$

For this choice of η and (28), we have

$$\begin{aligned} P_c^{(n)}(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)}) &\leq 5e^{-n\eta} \\ &= 5 \exp \left[-n[1 + \theta(2 + \alpha + \mu)]^{-1} \right. \\ &\quad \times \left. \left\{ \theta[R - \mu R_d] - \frac{1}{n} \Omega_{p^{(n)}||q^n}^{(\alpha, \mu, \theta)}(S^n X^n Y^n | K_n M_n) \right\} \right], \end{aligned}$$

completing the proof. \blacksquare

Set

$$\begin{aligned} \bar{\Omega}^{(\alpha, \mu, \theta)}(W) \\ \triangleq \sup_{n \geq 1} \max_{p^{(n)} \in \mathcal{P}^{(n)}(W)} \min_{q^n \in \mathcal{Q}^n} \frac{1}{n} \Omega_{p^{(n)}||q^n}^{(\alpha, \mu, \theta)}(S^n X^n Y^n | K_n M_n). \end{aligned}$$

Then we have the following corollary from Proposition 1.

Corollary 2: For any positive R, R_d and for any positive α, μ , and θ , we have

$$G(R_d, R|W) \geq \frac{\theta[R - \mu R_d] - \bar{\Omega}^{(\alpha, \mu, \theta)}(W)}{1 + \theta(2 + \alpha + \mu)}.$$

Proof: By the definition of $\bar{\Omega}^{(\alpha, \mu, \theta)}(W)$, the definition of $G^{(n)}(R_d, R|W)$, and Proposition 1, we have

$$G^{(n)}(R_d, R|W) \geq \frac{\theta[R - \mu R_d] - \bar{\Omega}^{(\alpha, \mu, \theta)}(W)}{1 + \theta(2 + \alpha + \mu)} - \frac{1}{n} \log 5,$$

from which we have Corollary 2. \blacksquare

We shall call $\bar{\Omega}^{(\alpha, \mu, \theta)}(W)$ the communication potential. The above corollary implies that the analysis of $\bar{\Omega}^{(\alpha, \mu, \theta)}(W)$ leads to an establishment of a strong converse theorem for the state dependent channels treated in this paper.

In the following argument we drive an explicit upper bound of $\bar{\Omega}^{(\alpha, \mu, \theta)}(W)$. We use a new technique we call the recursive method. The recursive method is a powerfull tool to drive a single letterized exponent function for rates below the rate distortion function. This method is also applicable to prove the exponential strong converse theorem for other network information theory problems [10], [11], [12].

For each $t = 1, 2, \dots, n$, define a function of $(u_t, v_t, s_t, x_t, y_t) \in \mathcal{U}_t \times \mathcal{V}_t \times \mathcal{S} \times \mathcal{X} \times \mathcal{Y}$ by

$$\begin{aligned} &f_{p_t||q_t}^{(\alpha, \mu, \theta)}(s_t, x_t, y_t|u_t, v_t) \\ &\triangleq \frac{W^{\alpha\theta}(y_t|x_t, s_t)}{q_{Y_t|X_t S_t U_t V_t}^{\alpha\theta}(y_t|x_t, s_t, u_t, v_t)} \\ &\quad \times \frac{q_{Y_t|U_t V_t}^\theta(y_t|u_t, v_t) q_{S_t|V_t}^\theta(s_t|v_t)}{q_{S_t|U_t V_t}^\theta(y_t|u_t, v_t) q_{Y_t|V_t}^\theta(y_t|v_t)} \frac{q_{Y_t|V_t}^\theta(y_t|v_t)}{q_{S_t|V_t}^\theta(s_t|v_t)} \frac{q_{S_t}^{\mu\theta}(s_t)}{q_{Y_t}^{\mu\theta}(y_t)} \\ &\quad \times \frac{p_{Y_t|U_t \hat{V}_t}^\theta(y_t|u_t, \hat{v}_t) p_{S_t|\hat{V}_t}^\theta(s_t|\hat{v}_t)}{q_{Y_t|U_t \hat{V}_t}^\theta(y_t|u_t, \hat{v}_t) q_{S_t|\hat{V}_t}^\theta(s_t|\hat{v}_t)} \frac{p_{S_t}^{\mu\theta}(s_t)}{q_{Y_t}^{\mu\theta}(y_t)} \frac{q_{Y_t}^{\mu\theta}(y_t)}{q_{Y_t|\hat{V}_t}^{\mu\theta}(y_t|\hat{v}_t)}. \end{aligned}$$

By definition we have

$$\begin{aligned} &\exp \left\{ \Omega_{p^{(n)}||q^n}^{(\alpha, \mu, \theta)}(S^n X^n Y^n | K_n M_n) \right\} \\ &= \sum_{k, s^n, x^n, y^n} p_{K_n S^n}(k, s^n) p_{X^n, Y^n | K_n S^n}(x^n, y^n | k, s^n) \\ &\quad \times \prod_{t=1}^n f_{p_t||q_t}^{(\alpha, \mu, \theta)}(s_t, x_t, y_t|u_t, v_t). \end{aligned} \quad (30)$$

For each $t = 1, 2, \dots, n$, we define a conditional probability distribution of (X^t, Y^t) given (K_n, S^n) by

$$\begin{aligned} &p_{X^t Y^t | K_n S^n}^{(\alpha, \mu, \theta; q^t)} \\ &\triangleq \left\{ p_{X^t Y^t | K_n S^n}^{(\alpha, \mu, \theta; q^t)}(x^t, y^t | k, s^n) \right\}_{(x^t, y^t, k, s^n) \in \mathcal{X}^t \times \mathcal{Y}^t \times \mathcal{K}_n \times \mathcal{S}^n}, \\ &p_{X^t Y^t | K_n S^n}^{(\alpha, \mu, \theta; q^t)}(x^t, y^t | k, s^n) \\ &\triangleq C_t^{-1}(k, s^n) p_{X^t Y^t | K_n S^n}(x^t, y^t | k, s^n) \\ &\quad \times \prod_{i=1}^t f_{p_i||q_i}^{(\alpha, \mu, \theta)}(s_i, x_i, y_i|u_i, v_i), \end{aligned}$$

where

$$\begin{aligned} C_t(k, s^n) &\triangleq \sum_{x^t, y^t} p_{X^t Y^t | K_n S^n}(x^t, y^t | k, s^n) \\ &\quad \times \prod_{i=1}^t f_{p_i||q_i}^{(\alpha, \mu, \theta)}(s_i, x_i, y_i|u_i, v_i) \end{aligned} \quad (31)$$

are constants for normalization. For $t = 1, 2, \dots, n$, define

$$\Phi_{t, q^t}^{(\alpha, \mu, \theta)}(k, s^n) \triangleq C_t(k, s^n) C_{t-1}^{-1}(k, s^n), \quad (32)$$

where we define $C_0(k, s^n) = 1$ for $(k, s^n) \in \mathcal{K}_n \times \mathcal{S}^n$. Then we have the following lemma.

Lemma 6: For each $t = 1, 2, \dots, n$, and for any $(k, s^n, x^t, y^t) \in \mathcal{K}_n \times \mathcal{S}^n \times \mathcal{X}^t \times \mathcal{Y}^t$, we have

$$\begin{aligned} &p_{X^t Y^t | K_n S^n}^{(\alpha, \mu, \theta; q^t)}(x^t, y^t | k, s^n) = (\Phi_{t, q^t}^{(\alpha, \mu, \theta)}(k, s^n))^{-1} \\ &\quad \times p_{X^{t-1} Y^{t-1} | K_n S^n}^{(\alpha, \mu, \theta; q^{t-1})}(x^{t-1}, y^{t-1} | k, s^n) \\ &\quad \times f_{p_t||q_t}^{(\alpha, \mu, \theta)}(s_t, x_t, y_t|u_t, v_t), \end{aligned} \quad (33)$$

$$\begin{aligned} &\Phi_{t, q^t}^{(\alpha, \mu, \theta)}(k, s^n) \\ &= \sum_{x^t, y^t} p_{X^{t-1} Y^{t-1} | K_n S^n}^{(\alpha, \mu, \theta; q^{t-1})}(x^{t-1}, y^{t-1} | k, s^n) \\ &\quad \times p_{X_t Y_t | K_n X^{t-1} S^n}(x_t, y_t | k, x^{t-1}, y^{t-1}, s^n) \\ &\quad \times f_{p_t||q_t}^{(\alpha, \mu, \theta)}(s_t, x_t, y_t|u_t, v_t). \end{aligned} \quad (34)$$

Furthermore, we have

$$\begin{aligned} &\exp \left\{ \Omega_{p^{(n)}||q^n}^{(\alpha, \mu, \theta)}(S^n X^n Y^n | K_n M_n) \right\} \\ &= \sum_{k, s^n} p_{K_n S^n}(k, s^n) \prod_{t=1}^n \Phi_{t, q^t}^{(\alpha, \mu, \theta)}(k, s^n). \end{aligned} \quad (35)$$

The equality (35) in Lemma 6 is obvious from (30), (31), and (32). Proofs of (33) and (34) in this lemma are given

in Appendix E. Next we define a probability distribution of random variable $K_n S^n$ taking values in $\mathcal{K}_n \times \mathcal{S}^n$ by

$$p_{K_n S^n}^{(\alpha, \mu, \theta; q^t)}(k, s^n) = \tilde{C}_t^{-1} p_{K_n S^n}(k, s^n) \prod_{i=1}^t \Phi_{i, q^i}^{(\alpha, \mu, \theta)}(k, s^n), \quad (36)$$

where \tilde{C}_t is a constant for normalization given by

$$\tilde{C}_t = \sum_{k, s^n} p_{K_n S^n}(k, s^n) \prod_{i=1}^t \Phi_{i, q^i}^{(\alpha, \mu, \theta)}(k, s^n).$$

For $t = 1, 2, \dots, n$, define

$$\Lambda_{t, q^t}^{(\alpha, \mu, \theta)} \triangleq \tilde{C}_t \tilde{C}_{t-1}^{-1}, \quad (37)$$

where we define $\tilde{C}_0 = 1$. Then we have the following.

Lemma 7:

$$\Omega_{p^{(n)} || q^n}^{(\alpha, \mu, \theta)}(S^n X^n Y^n | K_n M_n) = \sum_{t=1}^n \log \Lambda_{t, q^t}^{(\alpha, \mu, \theta)}, \quad (38)$$

$$\begin{aligned} \Lambda_{t, q^t}^{(\alpha, \mu, \theta)} &= \sum_{k, s^n} p_{K_n S^n}^{(\alpha, \mu, \theta; q^{t-1})}(k, s^n) \Phi_{t, q^t}^{(\alpha, \mu, \theta)}(k, s^n) \\ &= \sum_{k, s^n} \sum_{x^t, y^t} p_{K_n S^n}^{(\alpha, \mu, \theta; q^{t-1})}(k, s^n) \\ &\quad \times p_{X^{t-1} Y^{t-1} | K_n S^n}^{(\alpha, \mu, \theta; q^{t-1})}(x^{t-1}, y^{t-1} | k, s^n) \\ &\quad \times p_{X_t Y_t | X^{t-1} Y^{t-1} K_n S^n}(x_t, y_t | x^{t-1}, y^{t-1}, k, s^n) \\ &\quad \times f_{p_t || q_t}^{(\alpha, \mu, \theta)}(s_t, x_t, y_t | u_t, v_t). \end{aligned} \quad (39)$$

Proof: By the equality (35) in Lemma 6, we have

$$\begin{aligned} &\exp \left\{ \Omega_{p^{(n)} || q^n}^{(\alpha, \mu, \theta)}(S^n X^n Y^n | K_n M_n) \right\} \\ &= \tilde{C}_n = \prod_{t=1}^n \tilde{C}_t \tilde{C}_{t-1}^{-1} \stackrel{(a)}{=} \prod_{t=1}^n \Lambda_{t, q^t}^{(\alpha, \mu, \theta)}. \end{aligned} \quad (40)$$

Step (a) follows from the definition (37) of $\Lambda_{t, q^t}^{(\alpha, \mu, \theta)}$. From (40), we have (38) in Lemma 7. We next prove (39) in Lemma 7. Multiplying $\Lambda_{t, q^t}^{(\alpha, \mu, \theta)} = \tilde{C}_t / \tilde{C}_{t-1}$ to both sides of (36), we have

$$\begin{aligned} &\Lambda_{t, q^t}^{(\alpha, \mu, \theta)} p_{K_n S^n}^{(\alpha, \mu, \theta; q^t)}(k, s^n) \\ &= \tilde{C}_{t-1}^{-1} p_{K_n S^n}(k, s^n) \prod_{i=1}^t \Phi_{i, q^i}^{(\alpha, \mu, \theta)}(k, s^n), \\ &= p_{K_n S^n}^{(\alpha, \mu, \theta; q^{t-1})}(k, s^n) \Phi_{t, q^t}^{(\alpha, \mu, \theta)}(k, s^n). \end{aligned} \quad (41) \quad (42)$$

Taking summations of (41) and (42) with respect to (k, s^n) , we have (39) in Lemma 7. ■

The following proposition is a mathematical core to prove our main result.

Proposition 2: For $\theta \in (0, (2 + 2\mu)^{-1})$, set

$$\lambda = \frac{\theta}{1 - 2(1 + \mu)\theta} \Leftrightarrow \theta = \frac{\lambda}{1 + 2(1 + \mu)\lambda}. \quad (43)$$

Then, for any positive α, μ , and any $\theta \in (0, (2 + 2\mu)^{-1})$, we have

$$\bar{\Omega}^{(\alpha, \mu, \theta)}(W) \leq \frac{\Omega^{(\alpha, \mu, \lambda)}(W)}{1 + 2(1 + \mu)\lambda}.$$

Proof: Set

$$\begin{aligned} \hat{\mathcal{Q}}_n &\triangleq \{q = q_{UVSXY} : |\mathcal{U}| \leq |\mathcal{K}_n|, \\ &\quad |\mathcal{V}| \leq |\mathcal{M}_n| |\mathcal{S}^{n-1}| |\mathcal{Y}^{n-1}| \}, \\ \hat{\Omega}_n^{(\alpha, \mu, \lambda)}(W) &\triangleq \min_{q \in \hat{\mathcal{Q}}_n} \Omega_q^{(\alpha, \mu, \lambda)}(SXY | UV). \end{aligned}$$

Furthermore, set

$$\begin{aligned} &p_{K_n M_n S_t^n X_t Y_t}^{(\alpha, \mu, \theta; q^{t-1})}(k, m, s_t^n, x_t, y_t) \\ &= p_{U_t V_t S_t X_t Y_t}^{(\alpha, \mu, \theta; q^{t-1})}(u_t, v_t, s_t, x_t, y_t) \\ &\triangleq \sum_{x^{t-1}} \sum_{s^{t-1}, \phi^{(n)}(s^n)=m} p_{K_n S^n}^{(\alpha, \mu, \theta; q^{t-1})}(k, s^n) \\ &\quad \times p_{X^{t-1} Y^{t-1} | K_n S^n}^{(\alpha, \mu, \theta; q^{t-1})}(x^{t-1}, y^{t-1} | k, s^n) \\ &\quad \times p_{X_t Y_t | X^{t-1} Y^{t-1} K_n S^n}(x_t, y_t | x^{t-1}, y^{t-1}, k, s^n). \end{aligned} \quad (44)$$

Then by Lemma 7, we have

$$\begin{aligned} \Lambda_{t, q^t}^{(\alpha, \mu, \theta)} &= \sum_{u_t, v_t, s_t, x_t, y_t} p_{U_t V_t S_t X_t Y_t}^{(\alpha, \mu, \theta; q^{t-1})}(u_t, v_t, s_t, x_t, y_t) \\ &\quad \times f_{p_t || q_t}^{(\alpha, \mu, \theta)}(s_t, x_t, y_t | u_t, v_t). \end{aligned}$$

For each $t = 1, 2, \dots, n$, we choose $q_t = q_{U_t V_t S_t X_t Y_t}$ so that

$$q_{U_t V_t S_t X_t Y_t}(u_t, v_t, s_t, x_t, y_t) = p_{U_t V_t S_t X_t Y_t}^{(\alpha, \mu, \theta; q^{t-1})}(u_t, v_t, s_t, x_t, y_t)$$

and choose the following probability and conditional probability distributions:

$$\left. \begin{aligned} &q_{Y_t | X_t S_t U_t V_t}, q_{Y_t | U_t \hat{V}_t}, \\ &q_{Y_t | U_t V_t}, q_{Y_t | \hat{V}_t}, q_{Y_t}, \\ &q_{S_t | U_t V_t}, q_{S_t | \hat{V}_t}, q_{S_t} \end{aligned} \right\}$$

appearing in

$$\begin{aligned} &f_{p_t || q_t}^{(\alpha, \mu, \theta)}(s_t, x_t, y_t | u_t, v_t) \\ &= \frac{W^{\alpha\theta}(y_t | x_t, s_t)}{q_{Y_t | X_t S_t U_t V_t}^{\alpha\theta}(y_t | x_t, s_t, u_t, v_t)} \\ &\quad \times \frac{q_{Y_t | U_t V_t}^{\theta}(y_t | u_t, v_t) q_{S_t | V_t}^{\theta}(s_t | v_t) q_{Y_t | V_t}^{\mu\theta}(y_t | v_t) q_{S_t}^{\mu\theta}(s_t)}{q_{S_t | U_t V_t}^{\theta}(y_t | u_t, v_t) q_{Y_t | V_t}^{\theta}(y_t | v_t) q_{S_t | V_t}^{\mu\theta}(s_t | v_t) q_{Y_t}^{\mu\theta}(y_t)} \\ &\quad \times \frac{p_{Y_t | U_t \hat{V}_t}^{\theta}(y_t | u_t, \hat{v}_t) p_{S_t | \hat{V}_t}^{\theta}(s_t | \hat{v}_t) p_{S_t}^{\mu\theta}(s_t) q_{Y_t}^{\mu\theta}(y_t)}{q_{Y_t | U_t \hat{V}_t}^{\theta}(y_t | u_t, \hat{v}_t) q_{S_t | \hat{V}_t}^{\theta}(s_t | \hat{v}_t) q_{S_t}^{\mu\theta}(s_t) q_{Y_t | \hat{V}_t}^{\mu\theta}(y_t | \hat{v}_t)} \end{aligned}$$

such that they are the distributions induced by $q_{U_t V_t S_t X_t Y_t}$. Then for each $t = 1, 2, \dots, n$, we have the following chain of

inequalities:

$$\begin{aligned}
& \Lambda_{t,q^t}^{(\alpha,\mu,\theta)} \\
&= \mathbb{E}_{q_t} \left[\left\{ \frac{W^{\alpha\theta}(Y_t|X_t)}{q_{Y_t|X_t S_t U_t V_t}^{\alpha\theta}(Y_t|X_t, S_t, U_t, V_t)} \right. \right. \\
&\quad \times \frac{q_{Y_t|U_t V_t}^\theta(Y_t|U_t, V_t) q_{S_t|V_t}^\theta(S_t|V_t)}{q_{S_t|U_t V_t}^\theta(Y_t|U_t, V_t) q_{Y_t|V_t}^\theta(Y_t|V_t)} \\
&\quad \times \frac{q_{Y_t|V_t}^{\mu\theta}(Y_t|V_t) q_{S_t}^{\mu\theta}(S_t)}{q_{S_t|V_t}^{\mu\theta}(S_t|V_t) q_{Y_t}^{\mu\theta}(Y_t)} \left. \right\} \\
&\quad \times \left\{ \frac{p_{Y_t|U_t}^\theta(Y_t|U_t \hat{V}_t)}{q_{Y_t|U_t}^\theta(Y_t|U_t \hat{V}_t)} \right\} \left\{ \frac{p_{S_t|\hat{V}_t}^\theta(S_t|\hat{V}_t)}{q_{S_t|\hat{V}_t}^\theta(S_t|\hat{V}_t)} \right\} \\
&\quad \times \left\{ \frac{p_{S_t}^{\mu\theta}(S_t)}{q_{S_t}^{\mu\theta}(S_t)} \right\} \left\{ \frac{q_{Y_t}^{\mu\theta}(Y_t)}{q_{Y_t|\hat{V}_t}^{\mu\theta}(Y_t|\hat{V}_t)} \right\} \left. \right] \\
&\stackrel{(a)}{\leq} \left(\mathbb{E}_{q_t} \left[\left\{ \frac{W^{\alpha\theta}(Y_t|X_t)}{q_{Y_t|X_t S_t U_t V_t}^{\alpha\theta}(Y_t|X_t, S_t, U_t, V_t)} \right. \right. \right. \\
&\quad \times \frac{q_{Y_t|U_t V_t}^\theta(Y_t|U_t, V_t) q_{S_t|V_t}^\theta(S_t|V_t)}{q_{S_t|U_t V_t}^\theta(Y_t|U_t, V_t) q_{Y_t|V_t}^\theta(Y_t|V_t)} \\
&\quad \times \left. \left. \frac{q_{Y_t|V_t}^{\mu\theta}(Y_t|V_t) q_{S_t}^{\mu\theta}(S_t)}{q_{S_t|V_t}^{\mu\theta}(S_t|V_t) q_{Y_t}^{\mu\theta}(Y_t)} \right\}^{1-2(1+\mu)\theta} \right] \right)^{1-2(1+\mu)\theta} \\
&\quad \times \left\{ \mathbb{E}_{q_t} \left[\frac{p_{Y_t|U_t \hat{V}_t}(Y_t|U_t \hat{V}_t)}{q_{Y_t|U_t \hat{V}_t}(Y_t|U_t \hat{V}_t)} \right] \mathbb{E}_{q_t} \left[\frac{p_{S_t|\hat{V}_t}(S_t|\hat{V}_t)}{q_{S_t|\hat{V}_t}(S_t|\hat{V}_t)} \right] \right\}^\theta \\
&\quad \times \left\{ \mathbb{E}_{q_t} \left[\frac{p_{S_t}(S_t)}{q_{S_t}(S_t)} \right] \mathbb{E}_{q_t} \left[\frac{q_{Y_t}(Y_t)}{q_{Y_t|\hat{V}_t}(Y_t|\hat{V}_t)} \right] \right\}^{\mu\theta} \\
&= \exp \left\{ [1 - 2(1+\mu)\theta] \Omega_{q_t}^{(\alpha,\mu, \frac{\theta}{1-2(1+\mu)\theta})}(S_t X_t Y_t | U_t V_t) \right\} \\
&\stackrel{(b)}{=} \exp \left\{ \frac{\Omega_{q_t}^{(\alpha,\mu,\lambda)}(S_t X_t Y_t | U_t V_t)}{1 + 2(1+\mu)\lambda} \right\} \\
&\stackrel{(c)}{\leq} \exp \left\{ \frac{\hat{\Omega}_n^{(\alpha,\mu,\lambda)}(W)}{1 + 2(1+\mu)\lambda} \right\} \stackrel{(d)}{=} \exp \left\{ \frac{\Omega^{(\alpha,\mu,\lambda)}(W)}{1 + 2(1+\mu)\lambda} \right\}. \quad (45)
\end{aligned}$$

Step (a) follows from Hölder's inequality. Step (b) follows from (43). Step (c) follows from $q_t \in \hat{\mathcal{Q}}_n$ and the definition of $\hat{\Omega}_n^{(\alpha,\mu,\lambda)}(W)$. Step (d) follows from Lemma 10 in Appendix A. To prove this lemma we bound the cardinalities $|\mathcal{U}|$ and $|\mathcal{V}|$ appearing in the definition of $\hat{\Omega}_n^{(\alpha,\mu,\lambda)}(W)$ to show that the bounds $|\mathcal{U}|, |\mathcal{V}| \leq |\mathcal{S}| + |\mathcal{Y}| - 1$ are sufficient to describe $\hat{\Omega}_n^{(\alpha,\mu,\lambda)}(W)$. Hence we have the following:

$$\begin{aligned}
& \min_{q^n \in \mathcal{Q}^n} \frac{1}{n} \Omega_{p^{(n)}|q^n}^{(\alpha,\mu,\theta)}(S^n X^n Y^n | K_n M_n) \\
&\leq \frac{1}{n} \Omega_{p^{(n)}|q^n}^{(\alpha,\mu,\theta)}(S^n X^n Y^n | K_n M_n) \stackrel{(a)}{=} \frac{1}{n} \sum_{t=1}^n \log \Lambda_{t,q^t}^{(\alpha,\mu,\theta)} \\
&\stackrel{(b)}{\leq} \frac{\Omega^{(\alpha,\mu,\lambda)}(W)}{1 + 2(1+\mu)\lambda}. \quad (46)
\end{aligned}$$

Step (a) follows from (38) in Lemma 7. Step (b) follows from (45). Since (46) holds for any $n \geq 1$ and any $p^{(n)} \in \mathcal{P}^{(n)}$

(W), we have

$$\bar{\Omega}^{(\alpha,\mu,\theta)}(W) \leq \frac{\Omega^{(\alpha,\mu,\lambda)}(W)}{1 + 2(1+\mu)\lambda},$$

completing the proof. \blacksquare

Proof of Theorem 5: For $\theta \in (0, (2+2\mu)^{-1})$, set

$$\lambda = \frac{\theta}{1 - 2(1+\mu)\theta} \Leftrightarrow \theta = \frac{\lambda}{1 + 2(1+\mu)\lambda}. \quad (47)$$

Then we have the following:

$$\begin{aligned}
G(R_d, R|W) &\stackrel{(a)}{\geq} \frac{\theta(R - \mu R_d) - \bar{\Omega}^{(\alpha,\mu,\theta)}(W)}{1 + \theta(2 + \alpha + \mu)} \\
&\stackrel{(b)}{\geq} \frac{\lambda(R - \mu R_d) - \Omega^{(\alpha,\mu,\lambda)}(W)}{1 + 2(1+\mu)\lambda} \\
&\geq \frac{\lambda(2 + \alpha + \mu)}{1 + 2(1+\mu)\lambda} \\
&= \frac{\lambda(R - \mu R_d) - \Omega^{(\alpha,\mu,\lambda)}(W)}{1 + \lambda(4 + \alpha + 3\mu)} \\
&= F^{(\alpha,\mu,\lambda)}(R_d, R|W). \quad (48)
\end{aligned}$$

Step (a) follows from Corollary 2. Step (b) follows from Proposition 2 and (47). Since (48) holds for any positive α, μ , and λ , we have

$$G(R_d, R|W) \geq F(R_d, R|W).$$

Thus (3) in Theorem 5 is proved. The inclusion $\mathcal{R}(W) \subseteq \bar{\mathcal{R}}(W)$ is obvious from this bound. \blacksquare

IV. CONCLUSIONS

We have dealt with the state dependent discrete memoryless channels with full state information at the sender and partial state information at the receiver. We have proved that for rates outside the capacity region the correct probability of decoding tends to zero exponentially and derived an explicit lower bound of its exponent function.

APPENDIX

A. Cardinality Bound on Auxiliary Random Variables

Define

$$\begin{aligned}
\mathcal{Q}(W) &\triangleq \{q_{UVSXY} : |\mathcal{U}|, |\mathcal{V}| \leq |\mathcal{S}| + |\mathcal{Y}| - 1, \\
&\quad q_{Y|XS} = W, (U, V) \leftrightarrow (X, S) \leftrightarrow Y\}.
\end{aligned}$$

We first prove the following lemma.

Lemma 8:

$$\begin{aligned}
\bar{\mathcal{C}}^{(\mu)}(W) &\triangleq \max_{p \in \mathcal{P}(W)} \{I_p(U; Y|V) - I_p(U; S|V) \\
&\quad - \mu[I_p(V; S) - I_p(V; Y)]\} \\
&\leq \hat{\mathcal{C}}^{(\mu)}(W) \triangleq \max_{p \in \mathcal{Q}(W)} \{I_p(U; Y|V) - I_p(U; S|V) \\
&\quad - \mu[I_p(V; S) - I_p(V; Y)]\}
\end{aligned}$$

Proof: We first observe that

$$I_p(U; Y|V) - I_p(U; S|V) - \mu[I_p(V; S) - I_p(V; Y)] \\ = \sum_{v \in \mathcal{V}} p_V(v) \zeta_1^{(\alpha, \mu, \lambda)}(p_{USXY|V}(\cdot|v), p_S, p_Y), \quad (49)$$

where we set

$$\zeta_1^{(\alpha, \mu, \lambda)}(q_{USXY|V}(\cdot|v), p_S, p_Y) \\ \triangleq \sum_{(u, s, x, y) \in \mathcal{U} \times \mathcal{S} \times \mathcal{X} \times \mathcal{Y}} p_{USXY|V}(u, s, x, y|v) \\ \times \exp \left\{ \lambda \omega_p^{(\alpha, \mu)}(s, x, y|u, v) \right\}.$$

For each $v \in \mathcal{V}$, $\zeta_1^{(\alpha, \mu, \lambda)}(p_{USXY|V}(\cdot|v), p_S, p_Y)$ is a continuous function of $p_{USXY|V}(\cdot|v)$. We further observe that

$$\left. \begin{aligned} p_S(s) &= \sum_{v \in \mathcal{V}} q_V(v) q_{S|V}(s|v), \\ p_Y(y) &= \sum_{v \in \mathcal{V}} q_V(v) q_{Y|V}(y|v). \end{aligned} \right\} \quad (50)$$

Then by the support lemma,

$$|\mathcal{V}| \leq |\mathcal{S}| + |\mathcal{Y}| - 2 + 1 = |\mathcal{S}| + |\mathcal{Y}| - 1 \quad (51)$$

is sufficient to express $|\mathcal{S}| + |\mathcal{Y}| - 2$ values of (50) and one value of (49). We next bound the cardinality of \mathcal{U} on the conditional distribution $p_{U|V} = \{p_{U|V}(u|v)\}_{(u,v) \in \mathcal{U} \times \mathcal{V}}$. We first replace the set \mathcal{V} in (63) and (62) with $\tilde{\mathcal{V}}$ so that it satisfies the cardinality bound $|\tilde{\mathcal{V}}| \leq |\mathcal{S}| + |\mathcal{Y}| - 1$ and preserves the values of the right members of (50) and (49). Then we have

$$I_p(U; Y|V) - I_p(U; S|V) - \mu[I_p(V; S) - I_p(V; Y)] \\ = \sum_{v \in \tilde{\mathcal{V}}} p_V(v) \zeta_1^{(\alpha, \mu, \lambda)}(p_{USXY|V}(\cdot|v), p_S, p_Y). \quad (52)$$

Furthermore for each $v \in \tilde{\mathcal{V}}$, $\zeta_1^{(\mu)}(p_{USXY|V}(\cdot|v), p_S, p_Y)$ can be written as

$$\zeta_1^{(\alpha, \mu, \lambda)}(p_{USXY|V}(\cdot|v), p_S, p_Y) = \sum_{u \in \mathcal{U}} q_{U|V}(u|v) \\ \times \zeta_2^{(\alpha, \mu, \lambda)}(p_{SXY|UV}(\cdot|u, v), p_{S|V}(\cdot|v), p_{Y|V}(\cdot|v), p_S, q_Y). \quad (53)$$

Note that for each $v \in \tilde{\mathcal{V}}$, $\zeta_2^{(\alpha, \mu, \lambda)}$ is a continuous function of $q_{SXY|UV}(\cdot|u, v)$. For each $v \in \tilde{\mathcal{V}}$, we have

$$\left. \begin{aligned} p_{S|V}(s|v) &= \sum_{u \in \mathcal{U}} p_{U|V}(u|v) q_{S|UV}(s|u, v), \\ p_{Y|V}(y|v) &= \sum_{u \in \mathcal{U}} p_{U|V}(u|v) q_{Y|UV}(y|u, v). \end{aligned} \right\} \quad (54)$$

Then by the support lemma,

$$|\mathcal{U}| \leq |\mathcal{S}| + |\mathcal{Y}| - 2 + 1 = |\mathcal{S}| + |\mathcal{Y}| - 1 \quad (55)$$

is sufficient to express $|\mathcal{S}| + |\mathcal{Y}| - 2$ values of (67) and one value of (53). \blacksquare

We next prove the following lemma.

Lemma 9:

$$\hat{C}^{(\mu)}(W) \triangleq \max_{p \in \mathcal{Q}(W)} \{I_p(U; Y|V) - I_p(U; S|V) \\ - \mu[I_p(V; S) - I_p(V; Y)]\} \\ = C^{(\mu)}(W) \triangleq \max_{p \in \mathcal{P}_{\text{sh}}(W)} \{I_p(U; Y|V) - I_p(U; S|V) \\ - \mu[I_p(V; S) - I_p(V; Y)]\}.$$

Proof: We first bound the cardinality $|\mathcal{V}|$ of V to show that the bound $|\mathcal{V}| \leq |\mathcal{S}||\mathcal{X}|$ is sufficient to describe $\hat{C}^{(\mu)}(W)$. Observe that

$$I_p(U; Y|V) - I_p(U; S|V) - \mu[I_p(V; S) - I_p(V; Y)] \\ = \sum_{v \in \mathcal{V}} p_V(v) \zeta_1^{(\mu)}(p_{SX}, p_{USX|V}(\cdot|v)), \quad (56)$$

$$p_{SX}(s, x) = \sum_{v \in \mathcal{V}} p_V(v) p_{SX|V}(s, x|v), \quad (57)$$

where $\zeta_1^{(\mu)}$ is a continuous function of $p_{USX|V}(\cdot|v)$. Then by the support lemma,

$$|\mathcal{V}| \leq 1 + |\mathcal{S}||\mathcal{X}| - 1 = |\mathcal{S}||\mathcal{X}| \quad (58)$$

is sufficient to express one value of (57) and $|\mathcal{S}||\mathcal{X}| - 1$ values of (56). Next we derive an upper bound of $|\mathcal{U}|$. Observe that

$$I_p(U; Y|V) - I_p(U; S|V) - \mu[I_p(V; S) - I_p(V; Y)] \\ = \sum_{u \in \mathcal{U}} p_U(u) \zeta_2^{(\mu)}(p_{VSX}, p_{VSX|U}(\cdot|u)), \quad (59)$$

$$p_{VSX}(v, s, x) = \sum_{u \in \mathcal{U}} p_U(u) p_{VSX|U}(v, s, x|u), \quad (60)$$

where $\zeta_2^{(\mu)}$ is a continuous function of $p_{VSX|U}(\cdot|u)$. Then by the support lemma,

$$|\mathcal{U}| \leq 1 + |\mathcal{V}||\mathcal{S}||\mathcal{X}| - 1 = |\mathcal{V}||\mathcal{S}||\mathcal{X}| \quad (61)$$

is sufficient to express one value of (59) and $|\mathcal{V}||\mathcal{S}||\mathcal{X}| - 1$ values of (60). \blacksquare

Finally we prove the following lemma.

Lemma 10: For each integer $n \geq 2$, we define

$$\hat{\Omega}_n^{(\alpha, \mu, \lambda)}(W) \\ \triangleq \max_{\substack{q = q_{UVSXY}: \\ |\mathcal{V}| \leq |\mathcal{M}_n| |\mathcal{S}|^{n-1} |\mathcal{Y}|^{n-1}, \\ |\mathcal{U}| \leq |\mathcal{K}_n|}} \Omega_q^{(\alpha, \mu, \lambda)}(SXY|UV), \\ \Omega^{(\alpha, \mu, \lambda)}(W) \\ \triangleq \max_{\substack{q = q_{UVSXY}: \\ |\mathcal{U}|, |\mathcal{V}| \leq |\mathcal{S}| + |\mathcal{Y}| - 1}} \Omega_q^{(\alpha, \mu, \lambda)}(SXY|UV).$$

Then we have

$$\hat{\Omega}^{(\alpha, \mu, \lambda)}(W) = \Omega^{(\alpha, \mu, \lambda)}(W).$$

Proof: We first bound the cardinality $|\mathcal{V}|$ of V to show that the bound $|\mathcal{V}| \leq |\mathcal{S}| + |\mathcal{Y}| - 1$ is sufficient to describe $\hat{\Omega}_n^{(\alpha, \mu, \lambda)}(W)$. We first observe that

$$\Lambda_q^{(\alpha, \mu, \lambda)}(SXY|UV) \\ = \sum_{v \in \mathcal{V}} q_V(v) \zeta_3^{(\alpha, \mu, \lambda)}(q_{USXY|V}(\cdot|v), q_S, q_Y), \quad (62)$$

where we set

$$\begin{aligned} & \zeta_3^{(\alpha, \mu, \lambda)}(q_{USXY|V}(\cdot|v), q_S, q_Y) \\ & \triangleq \sum_{(u, s, x, y) \in \mathcal{U} \times \mathcal{S} \times \mathcal{X} \times \mathcal{Y}} q_{USXY|V}(u, s, x, y|v) \\ & \times \exp \left\{ \lambda \omega_q^{(\alpha, \mu)}(s, x, y|u, v) \right\}. \end{aligned}$$

For each $v \in \mathcal{V}$, $\zeta_3^{(\alpha, \mu, \lambda)}(q_{USXY|V}(\cdot|v), q_S, q_Y)$ is a continuous function of $q_{USXY|V}(\cdot|v)$. We further observe that

$$\left. \begin{aligned} q_S(s) &= \sum_{v \in \mathcal{V}} q_V(v) q_{S|V}(s|v), \\ q_Y(y) &= \sum_{v \in \mathcal{V}} q_V(v) q_{Y|V}(y|v). \end{aligned} \right\} \quad (63)$$

Then by the support lemma,

$$1 + |\mathcal{V}| \leq |\mathcal{S}| + |\mathcal{Y}| - 2 = |\mathcal{S}| + |\mathcal{Y}| - 1 \quad (64)$$

is sufficient to express one value of (62) and $|\mathcal{S}| + |\mathcal{Y}| - 2$ values of (63). We next bound the cardinality of \mathcal{U} on the conditional distribution $q_{U|V} = \{q_{U|V}(u|v)\}_{(u,v) \in \mathcal{U} \times \mathcal{V}}$. We first replace the set \mathcal{V} in (63) and (62) with $\tilde{\mathcal{V}}$ so that it satisfies the cardinality bound $|\tilde{\mathcal{V}}| \leq |\mathcal{S}| + |\mathcal{Y}| - 1$ and preserves the values of the right members of (63) and (62). Then we have

$$\begin{aligned} & \Lambda_q^{(\alpha, \mu, \lambda)}(SXY|UV) \\ &= \sum_{v \in \tilde{\mathcal{V}}} q_V(v) \zeta_3^{(\alpha, \mu, \lambda)}(q_{USXY|V}(\cdot|v), q_S, q_Y). \end{aligned} \quad (65)$$

Furthermore for each $v \in \tilde{\mathcal{V}}$, $\zeta_3^{(\mu)}(p_{USXY|V}(\cdot|v), q_S, q_Y)$ can be written as

$$\begin{aligned} & \zeta_3^{(\alpha, \mu, \lambda)}(q_{USXY|V}(\cdot|v), q_S, q_Y) = \sum_{u \in \mathcal{U}} q_{U|V}(u|v) \\ & \times \zeta_4^{(\alpha, \mu, \lambda)}(q_{SXY|UV}(\cdot|u, v), q_{S|V}(\cdot|v), q_{Y|V}(\cdot|v), q_S, q_Y). \end{aligned} \quad (66)$$

Note that for each $v \in \tilde{\mathcal{V}}$, $\zeta_4^{(\alpha, \mu, \lambda)}$ is a continuous function of $q_{SXY|UV}(\cdot|u, v)$. For each $v \in \tilde{\mathcal{V}}$, we have

$$\left. \begin{aligned} q_{S|V}(s|v) &= \sum_{u \in \mathcal{U}} q_{U|V}(u|v) q_{S|UV}(s|u, v), \\ q_{Y|V}(y|v) &= \sum_{u \in \mathcal{U}} q_{U|V}(u|v) q_{Y|UV}(y|u, v). \end{aligned} \right\} \quad (67)$$

Then by the support lemma,

$$|\mathcal{U}| \leq 1 + |\mathcal{S}| + |\mathcal{Y}| - 2 = |\mathcal{S}| + |\mathcal{Y}| - 1 \quad (68)$$

is sufficient to express one value of (66) and $|\mathcal{S}| + |\mathcal{Y}| - 2$ values of (67). ■

B. Proof of Property 1

In this appendix we prove Property 1. Property 1 part a) is a well known property. Proof of this property is omitted here. From Property 1 part a), we have the following lemma.

Lemma 11: Suppose that (\hat{R}_d, \hat{R}) does not belong to $\mathcal{C}(W)$. Then there exists $\epsilon, \mu^* > 0$ such that for any $(R_d, R) \in \mathcal{C}(W)$ we have

$$(R - \hat{R}) - \mu^*(R_d - \hat{R}_d) + \epsilon \leq 0.$$

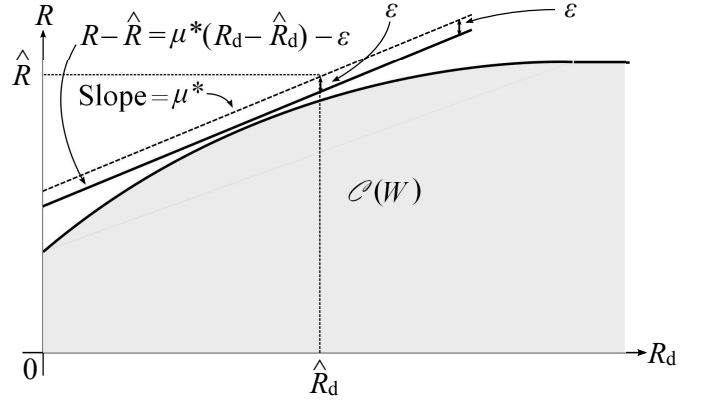


Fig. 4. An existence of a line that separates the point (\hat{R}_d, \hat{R}) from the region $\mathcal{C}(W)$.

Proof of this lemma is omitted here. Lemma 11 is equivalent to the fact that if the region $\mathcal{C}(W)$ is a convex set, then for any point (\hat{R}_d, \hat{R}) outside $\mathcal{C}(W)$, there exists a line which separates the point (\hat{R}_d, \hat{R}) from the region $\mathcal{C}(W)$. The line separating the point $(\hat{R}_d, \hat{R}) \notin \mathcal{C}(W)$ from the region $\mathcal{C}(W)$ is shown in Fig. 4. Lemma 11 will be used to prove Property 1 part b).

Proof of Property 1 part b): We first recall the following definitions of $\mathcal{P}(W)$ and $\mathcal{P}_{sh}(W)$:

$$\begin{aligned} \mathcal{P}(W) &\triangleq \{p_{UVSXY} : |\mathcal{V}| \leq |\mathcal{S}||\mathcal{X}| + 1, |\mathcal{U}| \leq |\mathcal{V}||\mathcal{S}||\mathcal{X}|, \\ & p_{Y|XS} = W, (U, V) \leftrightarrow (X, S) \leftrightarrow Y\}, \end{aligned}$$

$$\begin{aligned} \mathcal{P}_{sh}(W) &\triangleq \{p_{UVSXY} : |\mathcal{V}| \leq \min\{|\mathcal{S}||\mathcal{X}|, |\mathcal{S}| + |\mathcal{Y}| - 1\}, \\ & |\mathcal{U}| \leq \min\{|\mathcal{V}||\mathcal{S}||\mathcal{X}|, |\mathcal{S}| + |\mathcal{Y}| - 1\}, \\ & p_{Y|XS} = W, (U, V) \leftrightarrow (S, X) \leftrightarrow Y\}. \end{aligned}$$

We first prove $\mathcal{C}_{sh}(W) \subseteq \mathcal{C}(W)$. We assume that $(\hat{R}_d, \hat{R}) \notin \mathcal{C}(W)$. Then by Lemma 11, there exist $\epsilon, \mu^* > 0$ such that for any $(R_d, R) \in \mathcal{C}(W)$ we have

$$(R - \hat{R}) - \mu^*(R_d - \hat{R}_d) + \epsilon \leq 0.$$

Then we have

$$\begin{aligned} & \hat{R} - \mu^* \hat{R}_d \\ & \geq \max_{(R_d, R) \in \mathcal{C}(W)} \{R - \mu^* R_d\} + \epsilon \\ & \stackrel{(a)}{=} \max_{p \in \mathcal{P}(W)} \{I_p(U; Y|V) - I_p(U; S|V) \\ & \quad - \mu[I_p(V; S) - I_p(V; Y)]\} + \epsilon \\ & \geq \max_{p \in \mathcal{P}_{sh}(W)} \{I_p(U; Y|V) - I_p(U; S|V) \\ & \quad - \mu[I_p(V; S) - I_p(V; Y)]\} + \epsilon \\ & = \mathcal{C}^{(\mu^*)}(W) + \epsilon. \end{aligned} \quad (69)$$

Step (a) follows from the definition of $\mathcal{C}(W)$. The bound (69) implies that $(\hat{R}_d, \hat{R}) \notin \mathcal{C}_{sh}(W)$. Thus $\mathcal{C}_{sh}(W) \subseteq \mathcal{C}(W)$ is proved. We next prove $\mathcal{C}(W) \subseteq \mathcal{C}_{sh}(W)$. We assume that $(R_d, R) \in \mathcal{C}(W)$. Then there exists $p \in \mathcal{P}(W)$ such that

$$\left. \begin{aligned} R_d &\geq I_p(V; S) - I_p(V; Y), \\ R &\leq I_p(U; Y|V) - I_p(U; S|V). \end{aligned} \right\} \quad (70)$$

Define

$$\begin{aligned} \mathcal{Q}(W) &\triangleq \{q_{UVSXY} : |\mathcal{U}|, |\mathcal{V}| \leq |\mathcal{S}| + |\mathcal{Y}| - 1, \\ &\quad q_{Y|XS} = W, (U, V) \leftrightarrow (X, S) \leftrightarrow Y\}. \end{aligned}$$

Then, for $(R_d, R) \in \mathcal{C}(W)$, we have the following chain of inequalities:

$$\begin{aligned} R - \mu R_d &\stackrel{(a)}{\leq} I_p(U; Y|V) - I_p(U; S|V) - \mu[I_p(V; S) - I_p(V; Y)] \\ &\leq \max_{p \in \mathcal{P}(W)} \{I_p(U; Y|V) - I_p(U; S|V) \\ &\quad - \mu[I_p(V; S) - I_p(V; Y)]\} \\ &\stackrel{(b)}{\leq} \max_{p \in \mathcal{Q}(W)} \{I_p(U; Y|V) - I_p(U; S|V) \\ &\quad - \mu[I_p(V; S) - I_p(V; Y)]\} \\ &\stackrel{(c)}{=} \max_{p \in \mathcal{Q}(W)} \{-\alpha D(p_{Y|XSUV} \| W | p_{XSUV}) \\ &\quad + I_p(U; Y|V) - I_p(U; S|V) \\ &\quad - \mu[I_p(V; S) - I_p(V; Y)]\} \\ &\leq \max_{p \in \mathcal{Q}} \{-\alpha D(p_{Y|XSUV} \| W | p_{XSUV}) \\ &\quad + I_p(U; Y|V) - I_p(U; S|V) \\ &\quad - \mu[I_p(V; S) - I_p(V; Y)]\} \\ &= \tilde{C}^{(\alpha, \mu)}(W). \end{aligned}$$

Step (a) follows from (70). Step (b) follows from Lemma 8 stated in Appendix A. Step (c) follows from that when $p \in \mathcal{P}(W)$, we have

$$D(p_{Y|XSUV} \| W | p_{XSUV}) = D(p_{Y|XS} \| W | p_{XSUV}) = 0.$$

Hence we have $\mathcal{C}(W) \subseteq \tilde{\mathcal{C}}_{\text{sh}}(W)$. Finally we prove $\tilde{\mathcal{C}}_{\text{sh}}(W) \subseteq \mathcal{C}_{\text{sh}}(W)$. We assume that $(\tilde{R}_d, \tilde{R}) \in \tilde{\mathcal{C}}_{\text{sh}}(W)$. Then we have

$$\begin{aligned} \tilde{R} - \mu \tilde{R}_d &\leq \tilde{C}^{(\alpha, \mu)}(W) \\ &= \max_{q \in \mathcal{Q}} \{-\alpha D(q_{Y|XSUV} \| W | q_{XSUV}) \\ &\quad + I_q(U; Y|V) - I_q(U; S|V) \\ &\quad - \mu[I_q(V; S) - I_q(V; Y)]\} \\ &= -\alpha D(q_{Y|XSUV, \alpha, \mu}^* \| W | q_{XSUV, \alpha, \mu}^*) \\ &\quad + I_{q_{\alpha, \mu}^*}(U; Y|V) - I_{q_{\alpha, \mu}^*}(U; S|V) \\ &\quad - \mu[I_{q_{\alpha, \mu}^*}(V; S) - I_{q_{\alpha, \mu}^*}(V; Y)] \quad (71) \end{aligned}$$

where $q_{\alpha, \mu}^* = q_{UVSXY, \alpha, \mu}^* \in \mathcal{Q}$ is a probability distribution which attains the maximum in the definition of $\tilde{C}^{(\alpha, \mu)}(W)$. The quantities $q_{Y|XSUV, \alpha, \mu}^*$ and $q_{XSUV, \alpha, \mu}^*$ appearing in the first term in the right members of (71) is the conditional distributions induced by $q_{\alpha, \mu}^*$. We set

$$\begin{aligned} \Delta^{(\mu)} &\triangleq I_{q_{\alpha, \mu}^*}(U; Y|V) - I_{q_{\alpha, \mu}^*}(U; S|V) \\ &\quad - \mu[I_{q_{\alpha, \mu}^*}(V; S) - I_{q_{\alpha, \mu}^*}(V; Y)] - [\tilde{R} - \mu \tilde{R}_d] \end{aligned}$$

From (71), we must have

$$0 \leq \alpha D(q_{Y|XSUV, \alpha, \mu}^* \| W | q_{XSUV, \alpha, \mu}^*) \leq \Delta^{(\mu)}, \quad (72)$$

for any $\alpha, \mu > 0$. From (72), we have

$$0 \leq D(q_{Y|XSUV, \alpha, \mu}^* \| W | q_{XSUV, \alpha, \mu}^*) \leq \frac{\Delta^{(\mu)}}{\alpha}. \quad (73)$$

From (73), we have

$$\begin{aligned} \tilde{R} - \mu \tilde{R}_d &\leq I_{q_{\alpha, \mu}^*}(U; Y|V) - I_{q_{\alpha, \mu}^*}(U; S|V) \\ &\quad - \mu[I_{q_{\alpha, \mu}^*}(V; S) - I_{q_{\alpha, \mu}^*}(V; Y)] \quad (74) \end{aligned}$$

for any $\alpha, \mu > 0$. Let $\hat{q}_{\alpha, \mu} = \hat{q}_{UVSXY, \alpha, \mu}$ be a probability disitribution with the form

$$\hat{q}_{UVSXY, \alpha, \mu}(u, v, s, x, y) = q_{UVSX, \alpha, \mu}^*(u, v, s, x)W(y|x, s).$$

By definition, we have $\hat{q}_{\alpha, \mu} \in \mathcal{Q}(W)$. Computing $D(q_{\alpha, \mu}^* \| \hat{q}_{\alpha, \mu})$, we have the following:

$$\begin{aligned} D(q_{\alpha, \mu}^* \| \hat{q}_{\alpha, \mu}) &= D(q_{Y|XSUV, \alpha, \mu}^* \| W | q_{XSUV, \alpha, \mu}^*) \stackrel{(a)}{\leq} \frac{\Delta^{(\alpha, \mu)}}{\alpha}. \quad (75) \end{aligned}$$

Step (a) follows from (73). From (75), we have

$$\lim_{\alpha \rightarrow \infty} D(q_{\alpha, \mu}^* \| \hat{q}_{\alpha, \mu}) = 0,$$

from which we obtain

$$q_{\alpha, \mu}^* \rightarrow \hat{q}_{\alpha, \mu}, \quad \alpha \rightarrow \infty. \quad (76)$$

By (76) and the continuity of $I_q(U; Y|V)$, $I_q(U; S|V)$, $I_q(V; S)$, and $I_q(V; Y)$, with respect to the distribution $q = q_{UVSXY}$, we have that for any $\mu > 0$ and any sufficiently large α , we have

$$\begin{aligned} &I_{q_{\alpha, \mu}^*}(U; Y|V) - I_{q_{\alpha, \mu}^*}(U; S|V) \\ &\quad - \mu[I_{q_{\alpha, \mu}^*}(V; S) - I_{q_{\alpha, \mu}^*}(V; Y)] \\ &\leq I_{\hat{q}_{\alpha, \mu}}(U; Y|V) - I_{\hat{q}_{\alpha, \mu}}(U; S|V) \\ &\quad - \mu[I_{\hat{q}_{\alpha, \mu}}(V; S) - I_{\hat{q}_{\alpha, \mu}}(V; Y)] + \tau(\alpha, \mu), \quad (77) \end{aligned}$$

where $\tau(\alpha, \mu)$ is a positive number that satisfies

$$\lim_{\alpha \rightarrow \infty} \tau(\alpha, \mu) = 0.$$

Then we have the following chain of inequalities:

$$\begin{aligned} \tilde{R}_1 - \mu \tilde{R}_d &\stackrel{(a)}{\leq} I_{q_{\alpha, \mu}^*}(U; Y|V) - I_{q_{\alpha, \mu}^*}(U; S|V) \\ &\quad - \mu[I_{q_{\alpha, \mu}^*}(V; S) - I_{q_{\alpha, \mu}^*}(V; Y)] \\ &\stackrel{(b)}{\leq} I_{\hat{q}_{\alpha, \mu}}(U; Y|V) - I_{\hat{q}_{\alpha, \mu}}(U; S|V) \\ &\quad - \mu[I_{\hat{q}_{\alpha, \mu}}(V; S) - I_{\hat{q}_{\alpha, \mu}}(V; Y)] + \tau(\alpha, \mu) \\ &\stackrel{(c)}{\leq} \max_{q \in \mathcal{Q}(W)} \{I_q(U; Y|V) - I_q(U; S|V) \\ &\quad - \mu[I_q(V; S) - I_q(V; Y)]\} + \tau(\alpha, \mu) \\ &\stackrel{(d)}{=} \max_{q \in \mathcal{P}_{\text{sh}}(W)} \{I_q(U; Y|V) - I_q(U; S|V) \\ &\quad - \mu[I_q(V; S) - I_q(V; Y)]\} + \tau(\alpha, \mu) \\ &= C^{(\mu)}(W) + \tau(\alpha, \mu). \quad (78) \end{aligned}$$

Step (a) follows from (74). Step (b) follows from (77). Step (c) follows from that $\hat{q}_{\alpha, \mu} \in \mathcal{P}(p_{XY})$. Step (d) follows from that by Lemma 9 stated in Appendix A. Since in (78) the quantity $\tau(\alpha, \mu)$ can be made arbitrary close to zero, we conclude that $(\tilde{R}_d, \tilde{R}) \in \mathcal{C}_{\text{sh}}(W)$. Thus $\tilde{\mathcal{C}}_{\text{sh}}(W) \subseteq \mathcal{C}_{\text{sh}}(W)$ is proved. ■

C. Proof of Property 2

In this appendix we prove Property 2.

Proof of Property 2: We first prove parts a) and b). For simplicity of notations, set

$$\begin{aligned}\underline{a} &\triangleq (u, v, s, x, y), \underline{A} \triangleq (U, V, S, X, Y), \\ \underline{A} &\triangleq \mathcal{U} \times \mathcal{V} \times \mathcal{S} \times \mathcal{X} \times \mathcal{Y}, \\ \omega_q^{(\alpha, \mu, \lambda)}(s, x, y|u, v) &\triangleq \rho(\underline{a}), \\ \Omega_q^{(\alpha, \mu, \lambda)}(SXY|UV) &\triangleq \xi(\lambda).\end{aligned}$$

Then we have

$$\Omega_q^{(\alpha, \mu, \lambda)}(UVSXY) = \xi(\lambda) = \log \left[\sum_{\underline{a} \in \underline{A}} q_{\underline{A}}(\underline{a}) e^{\lambda \rho(\underline{a})} \right].$$

By simple computations we have

$$\begin{aligned}\xi'(\lambda) &= \left[\sum_{\underline{a}} q_{\underline{A}}(\underline{a}) e^{\lambda \rho(\underline{a})} \right]^{-1} \left[\sum_{\underline{a}} q_{\underline{A}}(\underline{a}) \rho(\underline{a}) e^{\lambda \rho(\underline{a})} \right], \\ \xi''(\lambda) &= \left[\sum_{\underline{a}} q_{\underline{A}}(\underline{a}) e^{\lambda \rho(\underline{a})} \right]^{-2} \\ &\times \left[\sum_{\underline{a}, \underline{b} \in \underline{A}} q_{\underline{A}}(\underline{a}) q_{\underline{A}}(\underline{b}) \frac{\{\rho(\underline{a}) - \rho(\underline{b})\}^2}{2} e^{\lambda \{\rho(\underline{a}) + \rho(\underline{b})\}} \right].\end{aligned}\quad (79)$$

$$\times \left[\sum_{\underline{a}, \underline{b} \in \underline{A}} q_{\underline{A}}(\underline{a}) q_{\underline{A}}(\underline{b}) \frac{\{\rho(\underline{a}) - \rho(\underline{b})\}^2}{2} e^{\lambda \{\rho(\underline{a}) + \rho(\underline{b})\}} \right]. \quad (80)$$

From (80), it is obvious that $\xi''(\lambda)$ is nonnegative. Hence $\Omega_q(SXY|UV)$ is a convex function of λ . From (79), we have

$$\begin{aligned}\xi'(0) &= \sum_{\underline{a}} q_{\underline{A}}(\underline{a}) \rho(\underline{a}) \\ &= -\alpha D(q_{Y|XSUV} || W | q_{XSUV}) \\ &\quad + I_q(U; Y|V) - I_q(U; S|V) \\ &\quad - \mu [I_q(V; S) - I_q(V; Y)].\end{aligned}\quad (81)$$

Hence we have the part b). Next we prove the part c). We assume that $(R, R_d, R) \notin \mathcal{C}(W)$, then by Property 1 part c), there exist positive α^* , μ^* , and ϵ such that

$$R - \mu^* R_d \geq \tilde{C}^{(\alpha^*, \mu^*)}(W) + \epsilon. \quad (82)$$

Set

$$\begin{aligned}\zeta(\lambda) &\triangleq \Omega_q^{(\alpha^*, \mu^*, \lambda)}(SXY|UV) \\ &\quad - \lambda \left[-\alpha^* D(q_{Y|XSUV} || W | q_{XSUV}) \right. \\ &\quad \left. + I_q(U; Y|V) - I_q(U; S|V) \right. \\ &\quad \left. - \mu^* [I_q(V; S) - I_q(V; Y)] + \frac{\epsilon}{2} \right].\end{aligned}$$

Then we have the following:

$$\begin{aligned}\zeta(0) &= 0, \zeta'(0) = -\frac{\epsilon}{2} < 0, \\ \zeta''(\lambda) &= \xi''(\lambda) \geq 0.\end{aligned}\quad (83)$$

It follows from (83) that there exists $\kappa(\epsilon) > 0$ such that we have $\zeta(\lambda) \leq 0$ for $\lambda \in (0, \kappa(\epsilon)]$. Hence for any $\lambda \in (0, \kappa(\epsilon)]$

and for every $q \in \mathcal{Q}$, we have

$$\begin{aligned}\Omega_q^{(\alpha^*, \mu^*, \lambda)}(SXY|UV) \\ \leq \lambda \left\{ -\alpha^* D(q_{Y|XSUV} || W | q_{XSUV}) \right. \\ \left. + I_q(U; Y|V) - I_q(U; S|V) \right. \\ \left. - \mu^* [I_q(V; S) - I_q(V; Y)] + \frac{\epsilon}{2} \right\}.\end{aligned}\quad (84)$$

From (84), we have that for any $\lambda \in (0, \kappa(\epsilon)]$,

$$\begin{aligned}\Omega^{(\alpha^*, \mu^*, \lambda)}(W) &= \max_{q \in \mathcal{Q}} \Omega_q^{(\alpha^*, \mu^*, \lambda)}(SXY|UV) \\ &\leq \lambda \max_{q \in \mathcal{Q}} \left\{ -\alpha^* D(q_{Y|XSUV} || W | q_{XSUV}) \right. \\ &\quad \left. + I_q(U; Y|V) - I_q(U; S|V) \right. \\ &\quad \left. - \mu^* [I_q(V; S) - I_q(V; Y)] + \frac{\epsilon}{2} \right\} \\ &= \lambda \left[\tilde{C}^{(\alpha^*, \mu^*)}(W) + \frac{\epsilon}{2} \right].\end{aligned}\quad (85)$$

Under (82) and (85), we have the following chain of inequalities:

$$\begin{aligned}F(R_d, R|W) &= \sup_{\alpha, \mu, \lambda > 0} F^{(\alpha, \mu, \lambda)}(R_d, R|W) \\ &\geq \sup_{\lambda \in (0, \kappa(\epsilon)]} F^{(\alpha^*, \mu^*, \lambda)}(R_d, R|W) \\ &= \sup_{\lambda \in (0, \kappa(\epsilon)]} \frac{\lambda(R - \mu^* R_d) - \Omega^{(\alpha^*, \mu^*, \lambda)}(W)}{1 + \lambda(4 + \alpha^* + 3\mu^*)} \\ &\stackrel{(a)}{\geq} \sup_{\lambda \in (0, \kappa(\epsilon)]} \frac{\lambda \left[(R - \mu^* R_d) - \tilde{C}^{(\alpha^*, \mu^*)}(W) - \frac{\epsilon}{2} \right]}{1 + \lambda(4 + \alpha^* + 3\mu^*)} \\ &\stackrel{(b)}{\geq} \sup_{\lambda \in (0, \kappa(\epsilon)]} \frac{1}{2} \cdot \frac{\lambda \epsilon}{1 + \lambda(4 + \alpha^* + 3\mu^*)} \\ &= \frac{1}{2} \cdot \frac{\kappa(\epsilon) \epsilon}{1 + \kappa(\epsilon)(4 + \alpha^* + 3\mu^*)} > 0.\end{aligned}$$

Step (a) follows from (85). Step (b) follows from (82). \blacksquare

D. Proof of Lemma 1

In this appendix we prove Lemma 1. For $(k, m) \in \mathcal{L}_n \times \mathcal{M}_n$, set

$$\begin{aligned}\mathcal{A}_1(k, m) &\triangleq \left\{ (s^n, x^n, y^n) : \right. \\ &\quad \left. \frac{1}{n} \log \frac{p_{Y^n|X^n S^n K_n M_n}(y^n | x^n, s^n, k, m)}{q_{Y^n|X^n S^n K_n M_n}^{(i)}(y^n | x^n, s^n, k, m)} \right. \\ &\quad \left. = \frac{1}{n} \log \frac{W(y^n | x^n, s^n)}{q_{Y^n|X^n S^n K_n M_n}^{(i)}(y^n | x^n, s^n, k, m)} \geq -\eta \right\}, \\ \tilde{\mathcal{A}}_2(k, m) &\triangleq \left\{ s^n : \frac{1}{n} \log \frac{p_{S^n|K_n M_n}(s^n | k, m)}{q_{S^n|K_n M_n}^{(ii)}(s^n | k, m)} \geq -\eta \right\}, \\ \mathcal{A}_2(k, m) &\triangleq \tilde{\mathcal{A}}_2(k, m) \times \mathcal{X}^n \times \mathcal{Y}^n.\end{aligned}$$

Furthermore, for $(k, m) \in \mathcal{K}_n \times \mathcal{M}_n$, set

$$\begin{aligned}\mathcal{A}_3(k, m) &\triangleq \{(s^n, x^n, y^n) : p_{Y^n|K_n M_n}(y^n|k, m) \\ &\quad \geq |\mathcal{K}_n| e^{-n\eta} q_{Y^n|M_n}^{(\text{iii})}(y^n|m)\}, \\ \mathcal{A}_4(k, m) &\triangleq \left\{ (s^n, x^n, y^n) : p_{S^n}(s^n) \right. \\ &\quad \left. \geq \frac{e^{-n\eta}}{|\mathcal{M}_n|} q_{Y^n|M_n}^{(\text{iv})}(y^n|m) \right\}, \\ \mathcal{A}(k, m) &\triangleq \bigcap_{i=1}^4 \mathcal{A}_i(k, m).\end{aligned}$$

Proof of Lemma 1: We have the following:

$$\begin{aligned}P_c^{(n)} &= \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{(s^n, x^n, y^n) \in \mathcal{A}(k, m), \\ y^n \in \mathcal{D}(k|m)}} \\ &\quad \times p_{K_n M_n S^n X^n Y^n}(k, m, s^n, x^n, y^n) \\ &\quad + \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{(s^n, x^n, y^n) \in \mathcal{A}^c(k, m): \\ y^n \in \mathcal{D}(k|m)}} \\ &\quad \times p_{K_n M_n S^n X^n Y^n}(k, m, s^n, x^n, y^n) \\ &\leq \sum_{i=0}^4 \Delta_i,\end{aligned}$$

where

$$\begin{aligned}\Delta_0 &\triangleq \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{(s^n, x^n, y^n) \in \mathcal{A}(k, m) \\ \times p_{K_n M_n S^n X^n Y^n}(k, m, s^n, x^n, y^n),}} \\ \Delta_i &\triangleq \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{(s^n, x^n, y^n) \in \mathcal{A}_i^c(k, m) \\ \times p_{K_n M_n S^n X^n Y^n}(k, m, s^n, x^n, y^n),}} \\ &\quad \text{for } i = 1, 2, 4, \\ \Delta_3 &\triangleq \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{(s^n, x^n, y^n) \in \mathcal{A}_3^c(k, m), \\ y^n \in \mathcal{D}(k|m)}} \\ &\quad \times p_{K_n M_n S^n X^n Y^n}(k, m, s^n, x^n, y^n) \\ &= \frac{1}{|\mathcal{K}_n|} \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{(s^n, x^n, y^n) \in \mathcal{A}_3^c(k, m), \\ y^n \in \mathcal{D}(k|m)}} \\ &\quad \times \varphi^{(n)}(x^n|k, s^n) W^n(y^n|x^n, s^n) p_{M_n S^n}(m, s^n).\end{aligned}$$

By definition we have

$$\begin{aligned}\Delta_0 &= p_{K_n M_n S^n X^n Y^n} \left\{ \right. \\ &\quad 0 \leq \frac{1}{n} \log \frac{W^n(Y^n|X^n S^n)}{q_{Y^n|X^n S^n K_n M_n}(Y^n|X^n, S^n, K_n, M_n)} + \eta, \\ &\quad 0 \leq \frac{1}{n} \log \frac{p_{S^n|K_n M_n}(S^n|K_n, M_n)}{q_{S^n|K_n M_n}^{(\text{ii})}(S^n|K_n, M_n)} + \eta, \\ &\quad \left. \frac{1}{n} \log |\mathcal{K}_n| \leq \log \frac{p_{Y^n|K_n M_n}(Y^n|K_n, M_n)}{q_{Y^n|M_n}^{(\text{iii})}(Y^n|M_n)} + \eta, \right\}\end{aligned}$$

$$\frac{1}{n} \log |\mathcal{M}_n| \geq \frac{1}{n} \log \frac{q_{S^n|M_n}^{(\text{iv})}(S^n|M_n)}{p_{S^n}(S^n)} - \eta \Big\}. \quad (86)$$

From (86), it follows that if $(\varphi^{(n)}, \phi^{(n)}, \psi^{(n)})$ satisfies

$$\frac{1}{n} \log |\mathcal{K}_n| \geq R, \quad \frac{1}{n} \log |\mathcal{M}_n| \leq R_d,$$

then the quantity Δ_0 is upper bounded by the first term in the right members of (8) in Lemma 1. Hence it suffices to show $\Delta_i \leq e^{-n\eta}$, $i = 1, 2, 3, 4$ to prove Lemma 1. We first prove $\Delta_i \leq e^{-n\eta}$ for $i = 1, 2$. We have the following chains of inequalities:

$$\begin{aligned}\Delta_1 &= \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{(s^n, x^n, y^n) \\ \in \mathcal{A}_1^c(k, m)}} p_{K_n M_n S^n X^n Y^n}(k, m, s^n, x^n, y^n) \\ &\leq e^{-n\eta} \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \\ &\quad \times \sum_{\substack{(s^n, x^n, y^n) \\ \in \mathcal{A}_1^c(k, m)}} q_{Y^n|X^n S^n K_n M_n}^{(\text{i})}(y^n|x^n, s^n, k, m) \\ &\quad \times p_{X^n S^n K_n M_n}(x^n, s^n, k, m) \\ &\leq e^{-n\eta}, \\ \Delta_2 &= \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{(s^n, x^n, y^n) \\ \in \mathcal{A}_2^c(k, m)}} p_{K_n M_n S^n X^n Y^n}(k, m, s^n, x^n, y^n) \\ &= \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{s^n \in \tilde{\mathcal{A}}_2^c(k, m)} p_{K_n M_n S^n}(k, m, s^n) \\ &\leq e^{-n\eta} \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \\ &\quad \times \sum_{s^n \in \tilde{\mathcal{A}}_2^c(k, m)} q_{S^n|K_n M_n}^{(\text{ii})}(s^n|k, m) p_{K_n M_n}(k, m) \\ &\leq e^{-n\eta}.\end{aligned}$$

Next, we prove $\Delta_3 \leq e^{-n\eta}$. We have the following chain of inequalities:

$$\begin{aligned}\Delta_3 &= \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{(s^n, x^n, y^n): \\ y^n \in \mathcal{D}(k|m), \\ p_{Y^n|K_n, M_n}(y^n|k, m) < e^{-n\eta} \\ \times |\mathcal{K}_n| q_{Y^n|M_n}^{(\text{iii})}(y^n|m)}} \\ &\quad \times p_{K_n M_n S^n X^n Y^n}(k, m, s^n, x^n, y^n) \\ &= \frac{1}{|\mathcal{K}_n|} \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{y^n \in \mathcal{D}(k|m), \\ p_{Y^n|K_n, M_n}(y^n|k, m) \\ < e^{-n\eta} |\mathcal{K}_n| q_{Y^n|M_n}^{(\text{iii})}(y^n|m)}} \\ &\quad \times p_{M_n}(m) p_{Y^n|K_n M_n}(y^n|k, m) \\ &\leq e^{-n\eta} \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{y^n \in \mathcal{D}(k|m)} p_{M_n}(m) q_{Y^n|M_n}^{(\text{iii})}(y^n|m) \\ &= e^{-n\eta} \sum_{m \in \mathcal{M}_n} p_{M_n}(m) q_{Y^n|M_n}^{(\text{iii})} \left(\bigcup_{k \in \mathcal{K}_n} \mathcal{D}(k|m) \middle| m \right) \\ &\leq e^{-n\eta} \sum_{m \in \mathcal{M}_n} p_{M_n}(m) = e^{-n\eta}.\end{aligned}$$

Finally, we prove $\Delta_4 \leq e^{-n\eta}$. We have the following chain of inequalities:

$$\begin{aligned}
\Delta_4 &= \sum_{(k,m) \in \mathcal{K}_n \times \mathcal{M}_n} \sum_{\substack{(s^n, x^n, y^n): \\ p_{S^n}(s^n) < \frac{e^{-n\eta}}{|\mathcal{L}_n|} \\ \times q_{S^n|M_n}^{(iv)}(s^n|m)}} \\
&\quad \times p_{K_n M_n S^n X^n Y^n}(k, m, s^n, x^n, y^n) \\
&= \sum_{m \in \mathcal{M}_n} \sum_{\substack{s^n: \phi^{(n)}(s^n)=m, \\ p_{S^n}(s^n) < \frac{e^{-n\eta}}{|\mathcal{M}_n|} \\ \times q_{S^n|M_n}^{(iv)}(s^n|m)}} p_{S^n}(s^n) \\
&\leq \frac{e^{-n\eta}}{|\mathcal{M}_n|} \sum_{m \in \mathcal{M}_n} \sum_{s^n: \phi^{(n)}(s^n)=m} q_{S^n|M_n}^{(iv)}(s^n|m) \\
&\leq \frac{e^{-n\eta}}{|\mathcal{M}_n|} \sum_{m \in \mathcal{M}_n} 1 = e^{-n\eta}.
\end{aligned}$$

Thus Lemma 1 is proved. \blacksquare

E. Proof of Lemma 6

In this appendix we prove (33) and (34) in Lemma 6.

Proofs of (33) and (34) in Lemma 6: By the definition of $p_{X^t Y^t | K_n S^n}^{(\alpha, \mu, \theta; q^t)}(x^t, y^t | k, s^n)$, for $t = 1, 2, \dots, n$, we have

$$\begin{aligned}
&p_{X^t Y^t | K_n S^n}^{(\alpha, \mu, \theta; q^t)}(x^t, y^t | k, s^n) \\
&= C_t^{-1}(k, s^n) p_{X^t Y^t | K_n S^n}(x^t, y^t | k, s^n) \\
&\quad \times \prod_{i=1}^t f_{p_i | q_i}^{(\alpha, \mu, \theta)}(s_i, x_i, y_i | u_i, v_i). \tag{87}
\end{aligned}$$

Then we have the following chain of equalities:

$$\begin{aligned}
&p_{X^t Y^t | K_n S^n}^{(\alpha, \mu, \theta; q^t)}(x^t, y^t | k, s^n) \\
&\stackrel{(a)}{=} C_t^{-1}(k, s^n) p_{X^t Y^t | K_n S^n}(x^t, y^t | k, s^n) \\
&\quad \times \prod_{i=1}^t f_{p_i | q_i}^{(\alpha, \mu, \theta)}(s_i, x_i, y_i | u_i, v_i) \\
&= C_t^{-1}(k, s^n) p_{X^{t-1} Y^{t-1} | K_n S^n}(x^{t-1}, y^{t-1} | k, s^n) \\
&\quad \times \prod_{i=1}^{t-1} f_{p_i | q_i}^{(\alpha, \mu, \theta)}(s_i, x_i, y_i | u_i, v_i) \\
&\quad \times p_{X_t Y_t | X^{t-1} Y^{t-1} K_n S^n}(x_t, y_t | x^{t-1}, y^{t-1}, k, s^n) \\
&\quad \times f_{p_t | q_t}^{(\alpha, \mu, \theta)}(s_t, x_t, y_t | u_t, v_t) \\
&\stackrel{(b)}{=} \frac{C_{t-1}(k, s^n)}{C_t(k, s^n)} p_{X^{t-1} Y^{t-1} | K_n S^n}^{(\alpha, \mu, \theta; q^{t-1})}(x^{t-1}, y^{t-1} | k, s^n) \\
&\quad \times p_{X_t Y_t | X^{t-1} Y^{t-1} K_n S^n}(x_t, y_t | x^{t-1}, y^{t-1}, k, s^n) \\
&\quad \times f_{p_t | q_t}^{(\alpha, \mu, \theta)}(s_t, x_t, y_t | u_t, v_t) \\
&= (\Phi_{t, q^t}^{(\alpha, \mu, \theta)}(k, s^n))^{-1} \\
&\quad \times p_{X^{t-1} Y^{t-1} | K_n S^n}^{(\alpha, \mu, \theta; q^{t-1})}(x^{t-1}, y^{t-1} | k, s^n) \\
&\quad \times p_{X_t Y_t | X^{t-1} Y^{t-1} K_n S^n}(x_t, y_t | x^{t-1}, y^{t-1}, k, s^n) \\
&\quad \times f_{p_t | q_t}^{(\alpha, \mu, \theta)}(s_t, x_t, y_t | u_t, v_t). \tag{88}
\end{aligned}$$

Steps (a) and (b) follow from (87). From (88), we have

$$\Phi_{t, q^t}^{(\alpha, \mu, \theta)}(k, s^n) p_{X^t Y^t | K_n S^n}^{(\alpha, \mu, \theta; q^t)}(x^t, y^t | k, s^n) \tag{89}$$

$$\begin{aligned}
&= p_{X^{t-1} Y^{t-1} | K_n S^n}^{(\alpha, \mu, \theta; q^{t-1})}(x^{t-1}, y^{t-1} | k, s^n) \\
&\quad \times p_{X_t Y_t | X^{t-1} Y^{t-1} K_n S^n}(x_t, y_t | x^{t-1}, y^{t-1}, k, s^n) \\
&\quad \times f_{p_t | q_t}^{(\alpha, \mu, \theta)}(s_t, x_t, y_t | u_t, v_t). \tag{90}
\end{aligned}$$

Taking summations of (89) and (90) with respect to x^t, y^t , we obtain

$$\begin{aligned}
&\Phi_{t, q^t}^{(\alpha, \mu, \theta)}(k, s^n) \\
&= \sum_{x^t, y^t} p_{X^{t-1} Y^{t-1} | K_n S^n}^{(\alpha, \mu, \theta; q^{t-1})}(x^{t-1}, y^{t-1} | k, s^n) \\
&\quad \times p_{X_t Y_t | X^{t-1} Y^{t-1} K_n S^n}(x_t, y_t | x^{t-1}, y^{t-1}, k, s^n) \\
&\quad \times f_{p_t | q_t}^{(\alpha, \mu, \theta)}(s_t, x_t, y_t | u_t, v_t),
\end{aligned}$$

completing the proof. \blacksquare

REFERENCES

- [1] S. Gel'fand and M. Pinsker, "Coding for channels with random parameters," *Prob. of Control and Information Theory*, vol.9, no.1, pp. 19-31, 1980.
- [2] H. Tyagi and P. Narayan, "The Gel'fand Pinsker channel: Strong converse and upper bound for the reliability function," *Proceedings of the IEEE International Symposium on Information Theory*, Seoul, Korea, pp.1954-1957, June 2009.
- [3] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Academic Press, New York, 1981.
- [4] C. Heegard and A. El Gamal, "On the capacity of computer memory with defect," *IEEE Trans. Inform Theory*, vol. IT-29, no. 5, pp.731-739, Sept. 1983.
- [5] Y. Steinberg, "Coding for channels with rate-limited side information at the decoder and applications," *IEEE Trans. Inform. Theory*, vol. 54, no. 9, pp. 4283-4295, Sept. 2008.
- [6] V. Y. F. Tan, "A formula for the capacity of the general Gel'fand-Pinsker channel," *Proceedings of the IEEE International Symposium on Information Theory*, Istanbul, Turkey, pp.2458-2462, July 2013.
- [7] T. S. Han, *Information-Spectrum Methods in Information Theory*. Springer-Verlag, Berlin, New York, 2002. The Japanese edition was published by Baifukan-publisher, Tokyo, 1998.
- [8] S. Arimoto, "On the converse to the coding theorem for discrete memoryless channels," *IEEE Trans. Inform. Theory*, vol. IT-19, no. 3, pp. 357-359, May 1973.
- [9] G. Dueck and J. Körner, "Reliability function of a discrete memoryless channel at rates above capacity," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 1, pp. 82-85, Jan. 1979.
- [10] Y. Oohama, "Exponent function for one helper source coding problem at rates outside the rate region," *Proceedings of the 2015 IEEE International Symposium on Information Theory*, pp. 1575-1579, Hong Kong, China, June 14-19, 2015.
- [11] Y. Oohama, "Strong converse exponent for degraded broadcast channels at rates outside the capacity region," *Proceedings of the 2015 IEEE International Symposium on Information Theory*, pp. 939 - 943, Hong Kong, China, June 14-19, 2015.
- [12] Y. Oohama, "Strong converse theorems for degraded broadcast channels with feedback," *Proceedings of the 2015 IEEE International Symposium on Information Theory*, pp. 2510 - 2514, Hong Kong, China, June 14-19, 2015.